

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2002, 21:44:06 ; Search time 1920.32 Seconds

(without alignments)
2741.112 Million cell updates/sec

Title: US-09-723-970-2

Perfect score: 390
Sequence: 1 atgaagctgcattccctcctt.....ccacgagctgtgagaccatc 390

Scoring table:

IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estio:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|-------|--------------------|---------|----------------------------|
| 1 | 286.2 | 73.4 | 680 9 | BB624070 BB624070 |
| 2 | 188 | 48.2 | 405 12 | AQ495547 HS-5211_B |
| 3 | 176.6 | 45.3 | 560 10 | BM262079 BM262079 |
| 4 | 171.2 | 43.9 | 617 9 | BE131943 BE131943 |
| 5 | 168.2 | 43.1 | 635 10 | BM262389 BM262389 |
| 6 | 72.6 | 18.6 | 928 12 | CNS0201E AL206123 Tetradon |
| 7 | 57.8 | 14.8 | 285 12 | BM327754 CH230-117 |
| 8 | 53 | 13.6 | 1595 11 | AK017593 Mus muscu |
| 9 | 50.8 | 13.0 | 592 12 | BM327754 CH230-108 |
| 10 | 49.8 | 12.8 | 423 10 | BM327754 CH230-108 |
| 11 | 49.8 | 12.8 | 423 10 | BM327754 CH230-108 |
| 12 | 49.4 | 12.7 | 497 9 | BE102709 UI-R-BT1- |
| 13 | 48.8 | 12.5 | 606 9 | BE1033915 QV2-NM200 |
| 14 | 48.2 | 12.4 | 372 10 | BE398865 UI-R-CAL- |
| 15 | 48 | 12.3 | 452 9 | AV746068 AV746068 |
| 16 | 47.8 | 12.3 | 268 12 | AZ871616 2M0184C002 |
| 17 | 47.8 | 12.3 | 522 9 | AV173269 AV173269 |

| | | | | |
|----|------|------|--------|---------------------|
| 18 | 47.8 | 12.3 | 523 9 | BB618381 BB618381 |
| 19 | 47.8 | 12.3 | 546 10 | BF074571 222022 MA |
| 20 | 47.8 | 12.3 | 584 10 | BM224915 K0203E03- |
| 21 | 47.4 | 12.2 | 584 10 | BM224971 K0204C05- |
| 22 | 47.2 | 12.1 | 453 9 | AI051683 OY7H08. x |
| 23 | 47.2 | 12.1 | 560 9 | AV752236 AV752236 |
| 24 | 47.2 | 12.1 | 678 9 | AV752236 AV752236 |
| 25 | 46.4 | 11.9 | 328 10 | BF079065 AV751658 |
| 26 | 46.4 | 11.9 | 379 9 | BF079065 AV751658 |
| 27 | 46.4 | 11.9 | 409 10 | BF442180 BF079065 |
| 28 | 46.4 | 11.9 | 444 10 | BF442180 BF079065 |
| 29 | 46.4 | 11.9 | 476 10 | BF075383 BF442180 |
| 30 | 46.4 | 11.9 | 482 10 | BF442020 BF075383 |
| 31 | 46.4 | 11.9 | 515 10 | BF080835 BF442020 |
| 32 | 46.4 | 11.9 | 516 10 | BF438666 BF080835 |
| 33 | 46.4 | 11.9 | 516 10 | BF438666 BF080835 |
| 34 | 46.4 | 11.9 | 527 10 | BF02350 BF438666 |
| 35 | 46.4 | 11.9 | 527 10 | BF02350 BF438666 |
| 36 | 45.8 | 11.7 | 624 9 | AV753384 AV753384 |
| 37 | 45.8 | 11.7 | 624 9 | AV753384 AV753384 |
| 38 | 45.4 | 11.6 | 481 9 | AV745740 AV745740 |
| 39 | 45.4 | 11.6 | 525 9 | AV745740 AV745740 |
| 40 | 44.8 | 11.5 | 544 9 | AI148647 qd02c08. x |
| 41 | 44.6 | 11.5 | 519 9 | AI148647 qd02c08. x |
| 42 | 44.6 | 11.4 | 447 9 | AI149163 qd02c08. x |
| 43 | 44 | 11.3 | 433 9 | AI291937 qm81c01. x |
| 44 | 44 | 11.3 | 487 10 | AV748886 AV748886 |
| 45 | 43.8 | 11.2 | 446 9 | BF190158 BF190158 |
| | | | | AI188964 AI188964 |
| | | | | AI276216 q176b05. x |

ALIGNMENTS

RESULT 1
LOCUS BB624070
DEFINITION BB624070 RIKEN full-length enriched, adult male eyeball Mus musculus cDNA clone 7530426119 5', mRNA sequence.
ACCESSION BB624070.1 GI:16462680
VERSION BB624070.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M., and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Wegli,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.

Db 82 AGACCAAAAGGTGACTGCAAGCTGCCAAGTGTGCCCGGAGTGCAGCCCTTCTACA 141
 QY 323 cctatccgctggccatccgctgtgactgagcgtgtccactgcccacacagagtgtrg 382
 Db 142 CCGATCCCGTGGCCATCCCTGTGACTGGAGCCCTGCTCCACTGCGACACAGAGTGTG 201
 QY 383 agacatc 390
 Db 202 AGACATC 209
 RESULT 3
 LOCUS BM262079/c 560 bp mRNA linear EST 18-DEC-2001
 DEFINITION dag40d11.x3 Blackshear/Soares normalized Xenopus egg library
 Xenopus laevis cDNA clone IMAGE:4783917 3' similar to SW:GTHB_CYPCA
 P01235 GONADOTROPIN BETA CHAIN PRECURSOR. [1] ; mRNA sequence.
 ACCESSION BM262079
 VERSION BM262079.1 GI:17925119
 KEYWORDS EST.
 ORGANISM African clawed frog.
 SOURCE Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 560)
 AUTHORS Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,
 Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
 ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
 Waterston,R. and Wilson,R.
 TITLE Washu Xenopus EST project, 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Sandy Clifton, Ph.D.
 Washu Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed by Bento Soares and M. Fatima Bonaldo
 (University of Iowa). DNA Sequencing by: Washington University
 Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LNL at: Info@image.lnl.gov
 Seq primer: -400P from GIBCO
 High quality sequence stop: 473.
 FEATURES
 source Location/Qualifiers
 1..560
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:4783917"
 /clone_lib="Blackshear/Soares normalized Xenopus egg
 library"
 /sex="female"
 /tissue_type="unfertilized egg"
 /cell_type="unfertilized egg"
 /dev_stage="unfertilized egg"
 /lab_host="DH10B"
 /note="Vector: pRT73-Pac; Site 1: EcoRI; Site 2: NotI;
 PolyA-selected mRNA was prepared from unfertilized Xenopus
 laevis eggs. The library was constructed in the vector
 pRT73-Pac as described in Bonaldo, M.F., Lennon, G. and
 Soares, M.B. 'Normalization and subtraction: two
 approaches to facilitate gene discovery', Genome Research
 6:791-806, 1996. The first strand synthesis used a
 NotI-drf8 primer: double stranded cDNAs were ligated to
 EcoRI adapters, digested with NotI, and directionally
 cloned into the NotI and EcoRI-digested pRT73-Pac vector.
 The library contained approximately 7.2 x 10⁵
 recombinants, with average insert sizes of 1-1.5 kb."
 BASE COUNT 164 a 112 c 120 g 164 t
 ORIGIN

Query Match 45.3%; Score 176.6; DB 10; Length 560;
 Best Local Similarity 73.0%; Pred. No. 1,1e-34;
 Matches 227; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
 QY 76 tccagtggaacatcgcaccccttctggcgtgtgcgtaaggaggttaacttctgcgc 135
 Db 543 TCCAAATATTAAGTGTGAGAACCTTATGGATGTGCTGTGAGGAATTCATTCTTACA 484
 QY 136 aagaagcgaagctgcaagggccttgatcacacagatgctgctgggtgcgtgag 195
 Db 483 AAGAAACCTGCTGCAAGAGTGTGCTGTGACTGATGATGCTGTGGGGGCGCTGAG 424
 QY 196 accctggagaacacccatcttgcgaaccccatatgaagcccatcagatcgtac 255
 Db 423 ACCGTGGAGAAAGCCATTCCTAGATCTCGGTATAGAGAACGCCACACAGATCTGCACT 364
 QY 256 tacaagagacccaacaggtgactgtcaagctgcgaactgtgcccgaggagtcaccc 315
 Db 363 TCAAAATGAACCTMAACTGTGTCTGTAACCTGCAACCTGCAGCCACGACATTCACCA 304
 QY 316 ttctaacctatcccgctgagcctcgcgtgtgactgagagcctgtccactgcacacag 375
 Db 303 TTCTTTACCTACCCAGTTGCCATTGATGTGACTGTGACATTGTTCACATTCTACTACA 244
 QY 376 gagtggagac 386
 Db 243 GAATGTGAGAC 233
 RESULT 4
 LOCUS BE131943 617 bp mRNA linear EST 29-JUN-2000
 DEFINITION db0d0d9.v1 Blackshear/Soares normalized xenopus egg library xenopus
 laevis cDNA clone IMAGE:3300977 5' similar to SW:GTHB_CYPCA P01235
 GONADOTROPIN BETA CHAIN PRECURSOR. [2] SW:GTHB_CYPCA ; mRNA
 sequence.
 ACCESSION BE131943
 VERSION BE131943.1 GI:8579306
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 617)
 AUTHORS Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,
 Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
 ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
 Waterston,R. and Wilson,R.
 TITLE Washu Xenopus EST project, 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Sandy Clifton, Ph.D.
 Washu Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed by Bento Soares and M. Fatima Bonaldo
 (University of Iowa). DNA Sequencing by: Washington University
 Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LNL at: Info@image.lnl.gov
 Seq primer: -400P from GIBCO
 High quality sequence stop: 467.
 FEATURES
 source Location/Qualifiers
 1..617
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:3300977"
 /clone_lib="Blackshear/Soares normalized Xenopus egg
 library"

BASE COUNT
ORIGIN

```

/sex="female"
/issue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10b"
/notes="Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI;
PolyA-selected mRNA was prepared from unfertilized Xenopus
laevis eggs. The library was constructed in the vector
pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalization and subtraction: two
approaches to facilitate gene discovery', Genome Research
6:791-806, 1996. The first strand synthesis used a
NotI-dt18 primer; double stranded cDNAs were ligated to
EcoRI adapters, digested with NotI, and directionally
cloned into the NotI and EcoRI-digested pT7T3-Pac vector.
The library contained approximately 7.2 x 105
recombinants, with average insert sizes of 1-1.5 kb."
```

189 a 139 c 151 g 138 t

| | | | | |
|---------------------------|-------|--------------------|-----------|-------------|
| Query Match | 43.9% | Score 171.2; | DB 9, | Length 617; |
| Best Local Similarity | 67.9% | Pred. No. 2.6e-33; | | |
| Matches 239; Conservative | 0; | Mismatches 113; | Indels 0; | Gaps 0; |

| | | | |
|----|-----|--|-----|
| Oy | 30 | catggccccccctcttccttgagctatgagctgctgcctcagtggaacct | 89 |
| Db | 266 | CCTGCAGCCTTCGTGGTTTAAAGCCCTGTGTCTACGACACCGTGCAGGATCCAAATATAGTCT | 325 |
| Oy | 90 | gcgcaccttltgaggctgtgcggtgaaggagttacttctctgccaagaagccagctg | 149 |
| Db | 336 | GAGAACTGTCATGGATGGATGTGCTGATGAGGGAATTCACATCTTAGGAAAGAAACCTGGCGTG | 385 |
| Oy | 150 | cagggagcctctggatcacaacacgatacgccgcgcggaggtctgtgagacctgggagaaac | 209 |
| Db | 386 | CAGAGTCTGGCGCTGTGCATCTGACATGATGAGCCCTGCTGGGGGCGCCTGTGAACTGTGAGAAAGCC | 445 |
| Oy | 210 | catcttgaaacccccctatatattgaagcccatcatcgagctctgtacctacaacagagaacaa | 269 |
| Db | 446 | ATCCCTAGATCCCTCCGTATCATATGAAAGCCACACACAGAGTCTCCACTTACAAATGAAACATTA | 505 |
| Oy | 270 | acaagttgactgtcaagctctcccacactgtgtcccgaggagtcgaaacctctacacatacc | 329 |
| Db | 506 | ACTGGTACTGTAAATACGTCCAAACGTGACGCCAACGACATTTGACCACTTTTAACTACACC | 565 |
| Oy | 330 | cgatggacatcgctgtgactgcggagacgtgtctccactgtccacacaaggagtg | 381 |
| Db | 566 | ACTTGCATTAATATGACTGTGACATGTGGTCCAACTTTCTACTACGAATGT | 617 |

RESULT 5
 BM262389
 LOCUS
 DEFINITION
 635 bp mRNA linear EST 18-DEC-2001
 accession
 BM262389
 version
 1
 keywords
 BM262389.1 GI:17925429
 EST.
 source
 African clawed frog.
 organism
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 635)
 Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.
 , Martin,J., Wylie,T., Underwood,R., Theisling,B., Bowers,Y., Person,
 B., Gibbons,M., Harvey,N., Rittler,E., Jackson,Y., McCann,R.,
 Waterson,R. and Wilson,R.
 Washu Xenopus EST project, 1999
 unpublished (1999)
 title
 JOURNAL
 COMMENT
 Other_ESTs: dag40d11.x3
 Contact: Sandy Clifton, Ph.D.

Mashu Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel.: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed by Bento Soares and M. Fatima Bonaldo
 (University of Iowa). DNA sequencing by Washington University
 Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LNL at: info@image.llnl.gov
 Seq primer: -40RP from Gibco
 High quality sequence stop: 491.
 Location/Qualifiers
 1. .635

```

187 a      153 C      153 g      142 T
BASE COUNT
ORIGIN
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="IMAGE:4783917"
/clone_lib="Blackshear/Soares normalized Xenopus egg
library"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"
/notes="Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI;
PolyA-selected mRNA was prepared from unfertilized Xenopus
laevis eggs. The library was constructed in the vector
pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalization and subtraction: two
approaches to facilitate gene discovery', Genome Research
6:791-806, 1996. The first strand synthesis used a
NotI-dn18 primer; double stranded cDNAs were ligated to
EcoRI adapters, digested with NotI, and directionally
cloned into the NotI and EcoRI-digested pT7T3-Pac vector.
The library contained approximately 7.2 X 105
recombinants, with average insert sizes of 1-1.5 kb."
```

| | | | | | | |
|----------|--|--------------|--------------------|-----------------|-----------------|---------|
| | Query Match | 43.1% | Score 168.2: | DB 10; | Length 635; | |
| | Best Local Similarity | 68.3%; | Pred. No. 1.5e-32; | | | |
| | Matches 233: | Conservative | 0; | Mismatches 108; | Indels 0; | Gaps 0; |
| Oy | 30 catggccctcctccttcgtgctgcgtatgctgctgtcctcggtgcctcagtgaggaaacct | 89 | | | | |
| | | | | | | |
| Db | 295 CCTGCAGCTTTCGGTTTATGCCCCCTGTGTCTCAGCACCGCTGCAGGATTCACATTAAAGTCT | 354 | | | | |
| Oy | 90 gcgcacatttltgaggctgtgccgtaaggagaattactttctcgtgccaagaagcaagctg | 149 | | | | |
| | | | | | | |
| Db | 355 GAGAAGTTCATTGGATGTGATGCTGTGTGAGGAATTCACATTCTTAGCAAAACCTCGCTG | 414 | | | | |
| Oy | 150 caggggccttcgtgatcaccaaggatgacctgcctcggggctgcttgtgagacttgggagaaaacc | 209 | | | | |
| | | | | | | |
| Db | 415 CAGAGTGTCGCTGTGACTACTGATGACCCTGTGGGGGCCCTGTGAACTGGGAGAAAGCC | 474 | | | | |
| Oy | 210 catctcggaaccccccatattggaagccatatcsgafctctactacaacgagaaacca | 269 | | | | |
| | | | | | | |
| Db | 475 ATNCCTAGATCTCTCCGTACTATAGMAAGCCACACAGAGCTGTGACTTACATAATGMAAAGCTAA | 534 | | | | |
| Oy | 270 acaagtlgactcgaagctctgcccacatctgccccgggagtcgaaccccttataccatlcac | 329 | | | | |
| | | | | | | |
| Db | 535 ACTGGTTACTGTAAAAACCTCCAAACTGCAGCCOACGACATTTGACOCCTTTTACTTAACCC | 594 | | | | |
| Oy | 330 cgltggccalcgcctgtgactcggagagcctgtctccactgcaca | 370 | | | | |
| | | | | | | |
| Db | 595 AGTTGGCATTAGATGTGACTGTGACATTGTGTTGCTACCTTCTA | 635 | | | | |
| RESULT 6 | CNS0201E/c | 928 bp | DNA | linear | GSS 14-MAY-2000 | |
| LOCUS | CNS0201E | | | | | |

DEFINITION Tetradodon nigroviridis genome survey sequence T7 end of clone 152J13 of library G from Tetradodon nigroviridis, genomic survey sequence.

ACCESSION AL206123.1 GI:7864942

VERSION AL206123

KEYWORDS GSS; genome survey sequence.

SOURCE Tetradodon nigroviridis.

ORGANISM Tetradodon nigroviridis

REFERENCE 1 (bases 1 to 928)

AUTHORS Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.

TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradodon nigroviridis

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 928)

AUTHORS Human gene number estimate provided by genome wide analysis using Tetradodon nigroviridis DNA sequence

TITLE Unpublished

JOURNAL Direct Submission

REFERENCE Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases

AUTHORS This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetradodon>.

FEATURES

source location/Qualifiers

1..928

/organism="Tetradodon nigroviridis"

/db_xref="taxon:99883"

/clone="152J13"

/clone_1lb="G"

/note="Genoscope sequence ID : COAG152CE07.LP1-end : T7"

BASE COUNT 266 a 214 c 238 g 196 t 14 others

ORIGIN

Query Match 18.6%; Score 72.6; DB 12; Length 928;

Best Local Similarity 77.0%; Pred. No. 3.3e-08;

Matches 87; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

QY 98 ttgtggagctgtgctgagagattacttctcgtgccaagaagcagctgcagggcc 157

DB 251 ttccggctgctggcggcgagcttctcgtgccaagaagcagctgcagggctc 192

QY 158 ttggatcaccaagagatgctgctgctgctgagacctggagaaacc 210

DB 191 tgcgcgtcacacagagcctgttggtggcgctgcacacacctggagagtaacc 139

RESULT 7

LOCUS BH327754 285 bp DNA linear GSS 03-DEC-2001

DEFINITION CH230-117A17.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone

ACCESSION BH327754

VERSION BH327754

KEYWORDS GSS.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 285)

AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn

TITLE 'A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de Jong, P. and Fraser, C.M.

JOURNAL Rat BAC End Sequences from library CHORI-230 EcoRI segment Unpublished (1999)

COMMENT Other_GSSs: CH230-117A17.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhaoe@tigr.org

Clones are derived from the rat BAC library CHORI-230 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library availability, please contact Pater de Jong (pdejong@tigr.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or_eirng_information.htm). BAC end page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html Plate: 117 row: A column: 17

Seq primer: T7

Class: BAC ends.

location/Qualifiers

1..285

/organism="Rattus norvegicus"

/strain="BN/SSNhd/MCw"

/db_xref="taxon:10116"

/clone="CH230-117A17"

/clone_1lb="CHORI-230 Segment 1"

/sex="Female"

/cell_type="Brain"

/note="Vector: PTARBAC2.1; Site.1: EcoRI; Site.2: EcoRI; CHORI-230 Rat (BN/SSNhd/MCw) BAC library produced by Pater de Jong"

BASE COUNT 60 a 66 c 83 g 76 t

ORIGIN

Query Match 14.8%; Score 57.8; DB 12; Length 285;

Best Local Similarity 96.7%; Pred. No. 0.00013;

Matches 59; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 144 agcgtcacaggagccttgatcaccaagatgctgctggtgctgtgagacctgga 203

DB 1 agcgtgcagggagcttgatcaccaagatgctgctggtgctgtgagacctgga 60

QY 204 g 204

DB 61 g 61

RESULT 8

LOCUS AK017593 1595 bp mRNA linear HTC 19-JAN-2002

DEFINITION AK017593

Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730420N04; follicle stimulating hormone beta, full insert sequence.

ACCESSION AK017593

VERSION AK017593

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (strain: C57BL/6J) 8 days embryo cDNA to mRNA, clone: 5730420N04.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (sites)

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2 (sites)

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,


```

/db_xref="taxon:3606"
/clone="TRPAEH04"
/clone_lib="np"
/tissue_type="pituitary tumor"
/dev_stage="Adult"
/lab_host="BM25.8"
/notes="Vector: pT7D3-Pac (Pharmacia) with a modified

BASE COUNT      156 a      143 c      156 g      149 t      2 others
ORIGIN

Query Match      12.5%; Score 48.8; DB 9; Length 606;
Best Local Similarity 60.6%; Pred. No. 0.031;
Matches 80; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

OY 230 tgcctcaacgagacacaacaggtgactgtgcaagctgcccgaactgtgccccggagatc 309
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 207 TGTACTCTCAAGAGACTGGTATATGAACAGAGAGAGCGCCGGCTGTGTCACCATGCA 266
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 310 gaccctcttaacctatcccggtgcaatccgctgtgactgtcgagagcctgtctcactgcc 369
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 267 GATTCTCTGTATACATACCCACAGTGGCCACCCAGTGTCACTGTGGCAAGTGTGACAGCAGC 326
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 370 accacgagatgtc 381
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 327 AGCACTGATGTGT 338

RESULT 14
LOCUS BF398865/c 372 bp mRNA Linear EST 27-NOV-2000
DEFINITION UT-R-CA1-b)a-a-15-0-UI.s1 UT-R-CA1 Rattus norvegicus cDNA clone
ACCESSION BF398865
VERSION BF398865
KEYWORDS BF398865.1 GI:11383867
SOURCE EST.
ORGANISM Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 372)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: mssoares@blue.weeg.iowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized corpus-striatum library cDNA library preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=yes.

FEATURES
SOURCE
1..372
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CA1-b)a-a-15-0-UI"
/clone_lib="UI-R-CA1"
/lab_host="MDH08 (Life Technologies)"
/notes="Vector: pT7D3-Pac (Pharmacia) with a modified

```

polylinker. site.1: Not I; site.2: Eco RI; The UI-R-CAL library is a subtracted library derived from the following tissues: thalamus, cerebellum, hypothalamus, medulla, pons, midbrain, cerebral cortex, corpus striatum, testis, and hippocampus. For a detailed description of the library from which this clone was derived, please visit our web site at ratel.eng.uiowa.edu. The subtraction has been previously described in (Bonaldi, Lennon and Soares, Genome Research 6:791-806, 1996)

TAG_LIB=UI-R-CAL
TAG_TISSUE=corpus-striatum
TAG_SEQ=CTAGC"

BASE COUNT 88 a 75 c 138 g 71 t
ORIGIN

Query Match 12.4%; Score 48.2; DB 10; Length 372;
Best Local Similarity 54.9%; Pred. No. 0.038;
Matches 95; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Oy 212 ttcttgaaacccccctatatltgaagccacalcagctctgtactacacagagacaaac 271
Db 332 TACTCCAGCTGCATGTGGCTCCCGGCTCAGACGAGTGTGACCTACGCTAGCTGGCGT 273

Oy 272 aggtgactgtcaagctgcccacactgtgtgcccgagtgatgcaccccttctaacctaccg 331
Db 272 TCGCCTCTGTGCGCGCTCCCTGCTCCGACCGCTGGGTAGACCCCACTGCTCTTCTCG 213

Oy 332 tggccatccgctgtgactgtcgagactctctcaacgcgacacagagatgtgag 384
Db 212 TGGCCCTCAGCTGCGCTGTGGGCCCTCGCTCAGTACGCTTACTGTGGG 160

RESULT 15
AV746068 452 bp mRNA linear EST 18-OCT-2000
LOCUS AV746068
DEFINITION AV746068 NPA Homo sapiens cDNA clone NPADEL10 5', mRNA sequence.
ACCESSION AV746068
VERSION AV746068.1 GI:10865515
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 452)
AUTHORS Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N.,
Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu
,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q.
, Han,Z., Chen,Z., Hu,R. and Chen,J.
Homo sapiens NPA library cDNA clones
Unpublished (2000)
Contact: Qinghua Zhang
Shanghai Institute of Endocrinology, Rui-jin Hospital
197 Rui-jin II Road, Shanghai 200025, P. R. China
Tel: 86-21-64370045(ex.663332)
Fax: 86-21-64743206
Email: mbsliems.stn.sh.cn
This clone is available at Shanghai Hematology Institute in
Shanghai.
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.
location/Qualifiers
1..452
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NPADEL10"
/clone_lib="NPA"
/tissue_type="pituitary"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 58 a 177 c 122 g 91 t 4 others

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2002, 21:21:20 ; Search time 1723.48 Seconds
(without alignments)
4735.390 Million cell updates/sec

Title: US-09-723-970-2

Perfect score: 390
Sequence: 1 atgaagctggcattctctt.....ccacgagtgtagacacac 390

Scoring table: IDENTITY_NTC
Gapox 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_pi:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pi:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_other:*
33: em_hgtg_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Length DB | ID | Description |
|------------|-------------|-----------|----|-------------|
|------------|-------------|-----------|----|-------------|

| | | | | | | |
|----|-------|-------|--------|---|------------|----------------------|
| 1 | 390 | 100.0 | 390 | 6 | AX259341 | AX259341 Sequence |
| 2 | 390 | 100.0 | 917 | 6 | AX201855 | AX201855 Sequence |
| 3 | 378 | 96.9 | 693 | 6 | AX193583 | AX193583 Sequence |
| 4 | 300.4 | 77.0 | 393 | 6 | AX259351 | AX259351 Sequence |
| 5 | 252 | 64.6 | 1045 | 9 | AX201857 | AX201857 Sequence |
| 6 | 204 | 52.3 | 12252 | 9 | CNS01DRS | AL118553 Human chr |
| 7 | 202 | 51.8 | 205 | 6 | AX193608 | AX193608 Sequence |
| 8 | 192 | 49.2 | 169650 | 9 | CNS00000 | AL049871 Human chr |
| 9 | 156 | 40.0 | 2985 | 6 | AX259362 | AX259362 Sequence |
| 10 | 156 | 17.7 | 1060 | 5 | ASPT51658 | AX251658 Actipenser |
| 11 | 66.4 | 17.0 | 687 | 5 | SCA310345 | AJ310345 Scylliorh |
| 12 | 61.6 | 15.8 | 436 | 6 | 108476 | 108476 Sequence 16 |
| 13 | 61.6 | 15.8 | 504 | 6 | CELHR | Y00518 Dog mRNA fo |
| 14 | 61.6 | 15.8 | 579 | 4 | AY056475 | AY056475 Monodelph |
| 15 | 61 | 15.6 | 426 | 6 | 108475 | 108475 Sequence 14 |
| 16 | 60.8 | 15.6 | 580 | 5 | ASPT51656 | AJ251656 Actipenser |
| 17 | 59.2 | 15.2 | 558 | 5 | CIGONADR | X61051 C. iodeila mr |
| 18 | 59.2 | 15.2 | 566 | 5 | AF319960 | AF319960 Mylophary |
| 19 | 58.4 | 15.0 | 429 | 4 | AF354938 | AF354938 Panthera |
| 20 | 58.4 | 15.0 | 448 | 5 | FUNCTHIT | M87015 Fundulus he |
| 21 | 57.8 | 14.8 | 525 | 6 | A06106 | A06106 Synthetic p |
| 22 | 57.8 | 14.8 | 629 | 4 | BOVLHBY | M10077 Bovine lutr |
| 23 | 57.4 | 14.7 | 2309 | 5 | AB015596 | AB015596 Carassius |
| 24 | 56.8 | 14.6 | 535 | 5 | D88024 | D88024 Carassius a |
| 25 | 56.8 | 14.6 | 5651 | 6 | AX211282 | AX211282 Sequence |
| 26 | 56.8 | 14.6 | 5651 | 6 | AX349366 | AX349366 Sequence |
| 27 | 56 | 14.4 | 426 | 6 | E01778 | E01778 DNA sequence |
| 28 | 56 | 14.4 | 581 | 5 | AF112192 | AF112192 Ictalurus |
| 29 | 56 | 14.4 | 661 | 5 | ONHSTHIB | M27154 Chum salmon |
| 30 | 56 | 14.4 | 999 | 5 | RI311355 | AJ311355 Rana ridl |
| 31 | 56 | 14.4 | 1864 | 4 | BOVLHBY | M11506 Bovine lute |
| 32 | 55.6 | 14.3 | 585 | 4 | AF090388 | AF090388 Trichosur |
| 33 | 55.6 | 14.3 | 601 | 4 | AF017448 | AF017448 Trichosur |
| 34 | 55.6 | 14.3 | 789 | 5 | AF112191 | AF112191 Ictalurus |
| 35 | 55.6 | 14.3 | 875 | 5 | ONHTRS | DI4652 Rainbow tro |
| 36 | 54.8 | 14.1 | 355 | 5 | AHA41769 | AJ41769 Atlantic |
| 37 | 54.6 | 14.0 | 527 | 5 | CHGTHIB | X91984 C. harengus |
| 38 | 54.6 | 14.0 | 533 | 5 | OALHBS | X52488 Sheep mRNA |
| 39 | 54.4 | 13.9 | 659 | 5 | AB050836 | AB050836 Oncorhyn |
| 40 | 54.4 | 13.9 | 1054 | 9 | AF397610 | AF397610 Cyprinopha |
| 41 | 54.2 | 13.9 | 590 | 5 | MOZSONATIB | L35066 Morone saxa |
| 42 | 54 | 13.8 | 810 | 5 | CCO271632 | AJ271632 Conger co |
| 43 | 54 | 13.8 | 866 | 5 | AF060566 | AF060566 Salmo sal |
| 44 | 54 | 13.8 | 1066 | 5 | AB016169 | AB016169 Anguilla |
| 45 | 53.6 | 13.7 | 426 | 4 | AF448455 | AF448455 Alluoropod |

ALIGNMENTS

RESULT 1
LOCUS AX259341 390 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 2 from Patent WO01/3034.
ACCESSION AX259341
VERSION AX259341.1 GI:16508575
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (sites)
AUTHORS Paszty,C.J., Cao,J., Danilenko,D.M., Gong,J. and Hill,D.C.
TITLE Beta-like glycoprotein hormone polypeptide and heterodimer
JOURNAL Patent: WO 01/3034-A 2 04-OCT-2001;
Amgen Inc. (US)
FEATURES
source
1..390
Location/Qualifiers
BASE COUNT 69 a 128 c 106 g 87 t
ORIGIN

[illegible]

| QY | 181 | tcggagtcgctcttggaagaccttggaagaataaaccaattcttggaacccccctatattgaagccat | 240 |
|--|---|---|--------|
| DB | 281 | tcggagtcgctcttggaagaccttggaagaataaaccaattcttggaacccccctatattgaagccat | 340 |
| QY | 241 | catcgaagctctgaccttaacaaacgaacccaagaagctgaaagctctgcaagcttgcccaactgctgc | 300 |
| DB | 341 | CATGAGTCTGTACTACTACAAAGACGACCAAAACAGGAGTGAAGTGTCAAGCTGCCAAGCTGTGCC | 400 |
| QY | 301 | ccggagatcgagaccctcttacaactatcccgctgagccatcgagctgtgagctcgagagccctgc | 360 |
| DB | 401 | CCGGAGTGCAGACCCCTTTTACACCTATCCGCTGGCCATCCGCTGTGACTGGGAGGACCTGC | 460 |
| QY | 361 | tcgaactgcaccacaaggagtgtagaacatc | 390 |
| DB | 461 | TCCACTGCCACCACGAGAGTGTGAGACCATC | 490 |
| RESULT | 3 | | |
| AX193583 | | | |
| LOCUS | AX193583 | 693 bp | DNA |
| DEFINITION | Sequence 5 from Patent WO0140291. | | Linear |
| ACCESSION | AX193583 | | |
| VERSION | AX193583.1 | GI:15211515 | |
| KEYWORDS | | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| AUTHORS | 1 (bases 1 to 693) | | |
| TITLE | Burgess,C.E., Prayaga,S.K., Shinkels,R.A., Raastelli,L., zerhusen,B.D., and Mezes,P.S. Proteins and nucleic acids encoding the same | | |
| JOURNAL | Patent: WO 0140291-A 5 07-JUN-2001; Curagen Corporation (US) | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..693 | | |
| BASE COUNT | 138 a 197 c 196 g 162 t | | |
| ORIGIN | | | |
| Query Match | 96.9%; Score 378; DB 6; Length 693; | | |
| Best Local Similarity | 100.0%; Pred. No. 5.1e-84; | | |
| Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| QY | 1 | atgaagctggaattcccttctccttgagcccatgagccctctcctctcttgctgagctatgagc | 60 |
| DB | 1 | ATGAAGCTGGGAGATTCTCTCTTGCCCGCCCAATGGCCCTCTCTTGCGTGGCGATATGGC | 60 |
| QY | 61 | tgatgccctcgctgcctcgaatggaagctgagcaacttgtagagctgtgacgtgagggag | 120 |
| DB | 61 | TGATGCTCTCGGTGCTCCAGTGGGAACTGGGCACCTTTGTGGGCTGTGCCGTAGGGAG | 120 |
| QY | 121 | ttacttctcctgagcaagaagccaagctgtagaggggaccttgatatacacaagatgacgc | 180 |
| DB | 121 | TTTACTTCTCTGGCCAAAGAGCAGGTGCAAGGGCTCTTCGGATCACCACGAGATGCTTGC | 180 |
| QY | 181 | tgaggctgctgtagaacttggaagaacccatctggaacccccctatatatgaagccat | 240 |
| DB | 181 | TGGGCTGCTGTGAGACCTGGGAGAAACCCATTCTGGAAACCCCTATATTGAAGCCAT | 240 |
| QY | 241 | catcgaagctctaacctaaacagagacaaacagatgagcttcaagctgccaactgtgac | 300 |
| DB | 241 | CATCAGTCTGTACTACTACAAAGACGACCAAAACGGGTGATGTCAAGCTGCCAAGCTGTGCC | 300 |
| QY | 301 | ccggagatcgagaccctcttacaactatcccgctgagccatcgagctgtgagctcgagagccctgc | 360 |
| DB | 301 | CCGGAGTGCAGACCCCTTTTACACCTATCCGCTGGCCATCCGCTGTGACTGGGAGGACCTGC | 360 |
| QY | 361 | tcgaactgcaccacaaggagtgtagaacatc | 390 |
| DB | 361 | TCCACTGCCACCACGAGAGTGTGAGACCATC | 390 |

Db 361 TCCACTGCCACGACGAG 378

RESULT 4
LOCUS AX259351 393 bp DNA Linear PAT 26-OCT-2001
DEFINITION Sequence 12 from Patent WO0173034.
ACCESSION AX259351
VERSION AX259351.1 GI:16508583
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
AUTHORS Paszty,C.J., Cao,J., Danilenko,D.M., Gong,J. and Hill,D.C.
TITLE Beta-like glycoprotein hormone polypeptide and heterodimer
JOURNML Patent: WO 0173034-A 12 04-OCT-2001;

FEATURES
Location/Qualifiers
1..393
/organism="Mus musculus"
/db_xref="taxon:10090"

BASE COUNT 78 a 111 c 109 g 95 t

ORIGIN

Query Match 77.0%; Score 300.4; DB 6; Length 393;
Best Local Similarity 85.6%; Pred. No. 1.2e-64;
Matches 334; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 atgaagctggcattctctctctctggcccatggccctctctctctggtggtatgac 60
Db 1 ATGAAGCTGGCATTCCTTCTTCTGTCAGTGGCCCTTCTCTCGGTGGCCCTGAC 60

QY 61 tgtgtccctcggtgctccagtggaactggcacccttgggtggtgtgtgctggagag 120
Db 61 TGTGTCTCGCATTCCTGAGTGGAACTGCACTTTGTGTGTGTGTGTGTGTGTGAGGAA 120

QY 121 ttacttctctggccaagaagccagctgcaaggcctctctctctctctctctctctgc 180
Db 121 TTCATCTTTCATGCGCAAGAAGCCAGGCTGCAAGGAGACTTGGATACCAAGATGCTTGC 180

QY 181 tgggtgcgtgtgagacctgggagaaccactctggaacccccctatatgtgaagccat 240
Db 181 TGGGTCGCTGTGAGACCTGGGAGAAACCACTGAGCCCTCCATATTGAAGCTTAT 240

QY 241 catgagctgtacctaaacagagcaaacaggtgactgttaagctgcccactgtgac 300
Db 241 CATGAGCTGTACCTAATGACACAGACAGGTGACAGTGAAGCTGCTAACTGTGTGC 300

QY 301 ccggagctgacccctctctacacctaccggtgacccgctgtgactgagctggagcctgc 360
Db 301 CCTGAGTGTGATACCTTTCTACACCTAACCTATAGGCTGTCCGATGTGAGTGTGGGCGTGT 360

QY 361 tccactgcccacagagagtggtgagaccatc 390
Db 361 TCCACTGCCACGACGAGTGTGAGACCATC 390

RESULT 5
AX201857 1045 bp DNA Linear PAT 30-AUG-2001
LOCUS AX201857
DEFINITION Sequence 3 from Patent WO0153346.
ACCESSION AX201857
VERSION AX201857.1 GI:15391697
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Mosselman,S. and spek van Der,P.J.

TITLE Human cystine knot polypeptide
JOURNML Patent: WO 0153346-A 3 26-JUL-2001;
AUTHORS Akzo Nobel N.V. (NL)
FEATURES
Source Location/Qualifiers
1..1045
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 254 a 279 c 234 g 278 t

ORIGIN

Query Match 64.6%; Score 252; DB 6; Length 1045;
Best Local Similarity 75.3%; Pred. No. 1.2e-52;
Matches 390; Conservative 0; Mismatches 0; Indels 128; Gaps 1;

QY 1 atgaagctggcattctctctctctggcccatggccctctctctctggtggtatgac 60
Db 101 ATGAAGCTGGCATTCCTTCTTCTGTCAGTGGCCCATGCGCCCTCTCTTGTGCTATGTC 160

QY 61 tgtgtccctcggtgctccagtggaactgcaaccttgggtggtgtgtgctggagagag 120
Db 61 TGTGTCTCGGTGCTCCAGTGGGAACCTGCGCACCTTTGTGTGTGTGTGTGTGTGTGAGGAG 220

QY 121 ttacttctctggccaagaagccagctgcaaggcctctctggtatcaccagagatgctgc 180
Db 221 TTTACTTCTCTGCGCCAAAGCCAGGCTGCAAGGGGCTTCTGATCCACGAGATGCTTGC 280

QY 181 tgggtgcgtgtgagacctggagagagagagagagagagagagagagagagagagag 204
Db 281 TGGGTGCTGTGTGAGACCTGGGAGGCTTTGTGTGATGATGCTGTGTGTGTGTGTGTGTG 340

QY 205 -----aaaccatcttggaaacccctatatgtg 204
Db 341 ATACATTTTGTGTGTGATGAGATGAGACCTCCCTCGAGCTGTATCTTCAGCT 400

QY 205 -----aaaccatcttggaaacccctatatgtg 232
Db 401 AATGAGAGCCATTGTGATGATGACACTTGTCACTTAACCACTTCTGTGAACTCTATATTG 460

QY 233 aagccatcatcgagctgtgtacctacacagagagcaagagtgtaactgaagctggcca 292
Db 461 AAGCCATCATGAGTGTGTACTTCTTCAACAGAGCAACCAAGTGTACTGACCTGCCCCA 520

QY 293 actgtgcccggagtgacacctctctacacctatcccggtgacatcccgctgtgactgcg 352
Db 521 ACTGTGCCCCGGAGTGTGACCCCTTCTACCTTATCCGTCGCAATCCGCTGTGACTGCG 580

QY 353 gagctgtctcaactgagccacagagagtgtagaccatc 380
Db 581 GAGCTGTCTCACTGCACTGCAACAGGAGTGTGAGACCATC 618

RESULT 6
CNS01DRS 122552 bp DNA Linear PRI 26-APR-2001
LOCUS CNS01DRS
DEFINITION Human chromosome 14 DNA sequence BAC C-2011M8 of library Caltech-P
ACCESSION AL118555
VERSION AL118555.6 GI:13513077
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Hellig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brothier,P., Calzavara,L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Eckenberg,R., Bruls,T., deBardina,V., Crouaud,C.,
Gyapay,G., Saurin,W. and Weissbach,J.
TITLE Unpublished
JOURNML Sequencing of the human chromosome 14
REFERENCE 2 (bases 1 to 122552)
AUTHORS Genoscope.

Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: <http://www.genoscope.cns.fr/>
Contact: SeqRef@genoscope.cns.fr

The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-111J20
Downstream BAC (overlapping the SP6 end) : C-2011M8 (AC-AL118555)

Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.69x in Q20 bases; sum-of-contigs

Overall quality chart :
Range : bases
0 :
1 - 9 :
10 - 19 :
20 - 29 :
30 - 39 :
40 - 49 :
50 - 59 :
60 - 69 :
70 - 79 :
80 - 89 :
90 - 99 :
Percentage of bases with a quality value >= 40 : 99 %

FEATURES

Location/Qualifiers

1..169650

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="14"

/clone="R-696D21"

/clone_lib="RPCI-11"

79560..79769

/note="matching EMBL:G11018"

RHdb:RH35164

dbSTS:STS8422

Identified using the e-PCR software (G. Schuler)"

123442..123577

RHdb:RH77820

dbSTS:STS4362

Identified using the e-PCR software (G. Schuler)"

124823..124920

/note="matching EMBL:AA011182"

RHdb:RH75585

dbSTS:STS2673

Identified using the e-PCR software (G. Schuler)"

125981..126128

/note="matching EMBL:N67960"

RHdb:RH65935

dbSTS:STS45865

Identified using the e-PCR software (G. Schuler)"

52841 a 35988 c 32593 g 48428 t

BASE COUNT

ORIGIN

FEATURES

Location/Qualifiers

1..169650

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="14"

/clone="R-696D21"

/clone_lib="RPCI-11"

79560..79769

/note="matching EMBL:G11018"

RHdb:RH35164

dbSTS:STS8422

Identified using the e-PCR software (G. Schuler)"

123442..123577

RHdb:RH77820

dbSTS:STS4362

Identified using the e-PCR software (G. Schuler)"

124823..124920

/note="matching EMBL:AA011182"

RHdb:RH75585

dbSTS:STS2673

Identified using the e-PCR software (G. Schuler)"

125981..126128

/note="matching EMBL:N67960"

RHdb:RH65935

dbSTS:STS45865

Identified using the e-PCR software (G. Schuler)"

52841 a 35988 c 32593 g 48428 t

BASE COUNT

ORIGIN

Query Match 49.2% Score 192; DB 9; Length 169650;
Best Local Similarity 99.5%; Pred. No. 7.3e-38;
Matches 203; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 atgaagctggatctctctctctggcccaatggccctctctctctgctgctatgac 60
Db 150808 ATGAAGCTGGATCTCTCTCTGGCCCAATGGCCCTCTCTCTGCTGCTATGAC 60
QY 61 tctgtccctcgtctcctcagtggaactgcgaaccttctgtggtcgtgctgagggag 120

|||||
Db 150748 TGTGCTCCTGCTGCTCCTCAGGAGGAGACCTGCGACCTTGTGGCTGTGCGTAGGGAG 150689
QY 121 tttcttctcctggcgaagaagcagctgagggccttctggtacacacagatgctgc 180
Db 150688 TTTCTCTCTCTGCGCAAGAGACCGAGCTGCGAGGG-CTTCGATACACAGATGCTGC 150630
QY 181 tggagctgctgagacctggag 204
Db 150629 TGGGCTGCTGTGAGACCTGCGAG 150606

RESULT 9

AX259362 2985 bp DNA Linear PAT 26-OCT-2001
LOCUS
DEFINITION Sequence 23 from Patent WO0173034.
ACCESSION AX259362
VERSION AX259362.1 GI:16508593
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (stiles)
Paszyk, C.J., Cao, J., Danilenko, D.M., Gong, J. and Hill, D.C.
Beta-like glycoprotein hormone polypeptide and heterodimer
Patent: WO 0173034-A 23 04-OCT-2001;
Amgen Inc. (US)

FEATURES

Location/Qualifiers

1..2985

/organism="Mus musculus"

/db_xref="taxon:10090"

BASE COUNT

ORIGIN

784 a 630 c 761 g 810 t

Query Match 40.0% Score 156; DB 6; Length 2985;
Best Local Similarity 85.3%; Pred. No. 9.6e-29;
Matches 174; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1 atgaagctggatctctctctctggcccaatggccctctctctctgctgctatgac 60
Db 1 ATGAAGCTGGATCTCTCTCTGGCCCAATGGCCCTCTCTCTGCTGCTATGAC 60

QY 61 tctgtccctcgtctcctcagtggaactgcgaaccttctggtcgtgctgagggag 120
Db 61 TCTGTCTCCTGAGCAGCTCAGTGGGAGACCTGTCACCTTTGTGGCTGTCTGTAGGGA 120

QY 121 ttacttctcctggcgaagaagcagctgacagggcctctggtacacacagatgctgc 180
Db 121 TTCACTTTCATGGCCCAAGAACCGCTGCGAGGGAGCTTGGATCACACAGATGCTGC 180

QY 181 tggagctgctgagacctggag 204
Db 181 TGGGCTGCTGTGAGACCTGCGAG 204

RESULT 10

AST251658 1060 bp mRNA Linear VRT 03-OCT-2000
LOCUS
DEFINITION Acipenser baeri mRNA for follicle-stimulating hormone (FSH gene).
ACCESSION AJ251658
VERSION AJ251658.1 GI:8250131
KEYWORDS follicle-stimulating hormone; FSH gene.
SOURCE Siberian sturgeon.
ORGANISM Acipenser baeri

REFERENCE
1 (bases 1 to 1060)
Queraf, B., Sellouk, A. and Salmon, C.
Phylogenetic analysis of the vertebrate glycoprotein hormone family
including new sequences of sturgeon (Acipenser baeri) beta subunits

OY 373 acgaggtgtgag 384
 |||||
 DB 368 TCAGACTGCGGG 379

RESULT 15
 108475 426 bp
 LOCUS Sequence 14 from Patent WO 8607383. linear PAT 02-DEC-1994

DEFINITION
 ACCESSION 108475
 VERSION 108475.1 GI:588818

KEYWORDS
 SOURCE
 ORGANISM

REFERENCE 1 (bases 1 to 426)
 AUTHORS Talmadge, K.D. and Fildes, J.C.
 TITLE AUTOANTIGEN VACCINES
 JOURNAL Patent: WO 8607383-A 14 18-DEC-1986;
 FEATURES Location/Qualifiers

source 1..426
 /organism="unknown"

BASE COUNT 56 a 161 c 126 g 83 t
 ORIGIN

Query Match 15.6%; Score 61; DB 6; Length 426;
 Best Local Similarity 52.6%; Pred. No. 5.3e-05;

Matches 133; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

OY 132 ggccaagaagcaggtcgaaggccctcgcgtacacacagatgcctcgtggtcgtcgtg 191
 |||||
 DB 111 GGCTGAGAGAGAGGCTGCCCTGCTGTATACACTTCACACACGACATCTGCGCGCTA 170
 OY 192 tgagacctgggagaacccatctcgaacccccctatatgaagccatcatcgtcgtcgt 251
 |||||
 DB 171 CTGCCCCACAGATGAGCGGCTGCTGCTGCTGCTGCCGCCGCGCGCGCGCGGCTG 230
 OY 252 tacctacaacgaagccaacagtgactgcaagctgcgaactgtgcacggtgcggtgagtcga 311
 |||||
 DB 231 CACTACCATGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 290
 OY 312 cccctctacaacctaccggtgcaatcgcgtgactgcgagactgcgtgcactgcacac 371
 |||||
 DB 291 CCCAATGCT 350
 OY 372 caagaggtgtgag 384
 |||||
 DB 351 CACTGACTGTGGG 363

Search completed: July 8, 2002, 13:28:38
 Job time: 58038 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2002, 21:20:08 ; Search time 238.01 Seconds
(without alignments)
2813.313 Million cell updates/sec

Title: US-09-723-970-2

Perfect score: 390

Sequence: 1 atgaagctgcatctctc.....ccacgagctgagacacac 390

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 390 | 100.0 | 390 | 22 | AAH46586 |
| 2 | 390 | 100.0 | 390 | 24 | AAH46586 |
| 3 | 390 | 100.0 | 390 | 24 | AAH46586 |
| 4 | 390 | 100.0 | 390 | 24 | AAH46586 |
| 5 | 378 | 96.9 | 390 | 22 | AAH46586 |
| 6 | 318 | 81.5 | 390 | 22 | AAH46586 |
| 7 | 300.4 | 77.0 | 390 | 24 | AAH46586 |
| 8 | 290.8 | 74.6 | 390 | 24 | AAH46586 |
| 9 | 260 | 66.7 | 318 | 22 | AAH46586 |

| | | | | | |
|----|------|------|------|----|----------|
| 10 | 252 | 64.6 | 1045 | 22 | AAH42568 |
| 11 | 156 | 40.0 | 2985 | 24 | AAH42568 |
| 12 | 61.6 | 15.8 | 496 | 7 | AAH42568 |
| 13 | 57.8 | 14.8 | 525 | 21 | AAH42568 |
| 14 | 56.8 | 14.6 | 535 | 22 | AAH42568 |
| 15 | 56.8 | 14.6 | 5651 | 22 | AAH42568 |
| 16 | 53 | 13.6 | 524 | 19 | AAH42568 |
| 17 | 53 | 13.6 | 611 | 11 | AAH42568 |
| 18 | 51.4 | 13.2 | 843 | 19 | AAH42568 |
| 19 | 50.4 | 12.9 | 616 | 11 | AAH42568 |
| 20 | 50.2 | 12.9 | 486 | 10 | AAH42568 |
| 21 | 48.8 | 12.5 | 278 | 21 | AAH42568 |
| 22 | 48.8 | 12.5 | 309 | 21 | AAH42568 |
| 23 | 48.8 | 12.5 | 312 | 21 | AAH42568 |
| 24 | 48.8 | 12.5 | 312 | 21 | AAH42568 |
| 25 | 48.8 | 12.5 | 315 | 21 | AAH42568 |
| 26 | 48.8 | 12.5 | 315 | 21 | AAH42568 |
| 27 | 48.8 | 12.5 | 315 | 21 | AAH42568 |
| 28 | 48.8 | 12.5 | 318 | 21 | AAH42568 |
| 29 | 48.8 | 12.5 | 318 | 21 | AAH42568 |
| 30 | 48.8 | 12.5 | 318 | 21 | AAH42568 |
| 31 | 48.8 | 12.5 | 321 | 21 | AAH42568 |
| 32 | 48.8 | 12.5 | 321 | 21 | AAH42568 |
| 33 | 48.8 | 12.5 | 321 | 21 | AAH42568 |
| 34 | 48.8 | 12.5 | 324 | 21 | AAH42568 |
| 35 | 48.8 | 12.5 | 324 | 21 | AAH42568 |
| 36 | 48.8 | 12.5 | 324 | 21 | AAH42568 |
| 37 | 48.8 | 12.5 | 327 | 21 | AAH42568 |
| 38 | 48.8 | 12.5 | 327 | 21 | AAH42568 |
| 39 | 48.8 | 12.5 | 327 | 21 | AAH42568 |
| 40 | 48.8 | 12.5 | 330 | 21 | AAH42568 |
| 41 | 48.8 | 12.5 | 330 | 21 | AAH42568 |
| 42 | 48.8 | 12.5 | 333 | 21 | AAH42568 |
| 43 | 48.8 | 12.5 | 333 | 21 | AAH42568 |
| 44 | 48.8 | 12.5 | 348 | 18 | AAH42568 |
| 45 | 48.8 | 12.5 | 348 | 18 | AAH42568 |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AAH46586 | AAH46586 standard; DNA; 390 BP. |
| XX | XX |
| AC | AAH46586; |
| XX | XX |
| DT | 17-SEP-2001 (first entry) |
| XX | XX |
| DE | Human anterior pituitary hormone-related polynucleotide #1. |
| XX | XX |
| KW | Human: anterior pituitary hormone; hypertension; autoimmune disease; |
| KW | heart failure; ds. |
| XX | XX |
| OS | Homo sapiens. |
| XX | XX |
| PN | W0200144475-A1. |
| XX | XX |
| PD | 21-JUN-2001. |
| XX | XX |
| PF | 15-DEC-2000; 2000MO-JP08896. |
| XX | XX |
| PR | 17-DEC-1999; 99JP-0358707. |
| XX | XX |
| PR | 18-FEB-2000; 2000JP-0046825. |
| XX | XX |
| PA | (TAKE) TAKEDA CHEM IND LTD. |
| XX | XX |
| PI | Hinuma S, Fukusumi S, Fujii R, Hosoya M; |
| XX | XX |
| DR | WPI: 2001-408485/43. |
| XX | XX |
| DR | P-PDB; AAC64064. |
| XX | XX |
| PT | Polypeptides for treatment of hypertension, autoimmune disease and |

```
PT heart failure -
XX
XX Claim 7, Fig 1; 107pp; Japanese.
XX
XX The invention relates to a novel polypeptide comprising a fully defined
CC 130 amino acid sequence given in the specification and its amides,
CC esters and salts. The polypeptide has anterior pituitary hormone-related
CC activity. It is useful for the treatment of hypertension, autoimmune
CC diseases and heart failure. The screening method and kit also
CC provided in the invention are useful for identifying new substances
CC for treating and preventing these diseases. The present sequence encodes
CC the polypeptide of the invention.
XX
XX Sequence 390 BP; 69 A; 128 C; 106 G; 87 T; 0 other;

Query Match 100.0%; Score 390; DB 22; Length 390;
Best Local Similarity 100.0%; Pred. No. 4,7e-102;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgaagctggcattctctctctctggcccatgagccctcctctctgctgctatgac 60
    |||
DB 1 atgaagctggcattctctctctctggcccatgagccctcctctctgctgctatgac 60
    |||

QY 61 tgtgtcctcggtgctccagtggaacctgagcacttgggtggtcgtgcgtgaaggag 120
    |||
DB 61 tgtgtcctcggtgctccagtggaacctgagcacttgggtggtcgtgcgtgaaggag 120
    |||

QY 121 ttactctctggccaagaagcagctgcaaggccttcgataccaacagatgctgc 180
    |||
DB 121 ttactctctggccaagaagcagctgcaaggccttcgataccaacagatgctgc 180
    |||

QY 181 tgggtcgtcgtgtagagactgggaaccattctggaacccctatatattgaagccat 240
    |||
DB 181 tgggtcgtcgtgtagagactgggaaccattctggaacccctatatattgaagccat 240
    |||

QY 241 catcaggtctgtacctaccacaagcagagctgacttcaagctggcccaactgtgc 300
    |||
DB 241 catcaggtctgtacctaccacaagcagagctgacttcaagctggcccaactgtgc 300
    |||

QY 301 ccggaggtcgaccctctctacactatccgttggccatccgctgtgactcggaagcctgc 360
    |||
DB 301 ccggaggtcgaccctctctacactatccgttggccatccgctgtgactcggaagcctgc 360
    |||

QY 361 tccactgcccacacagagatgtgagaccatc 390
    |||
DB 361 tccactgcccacacagagatgtgagaccatc 390
    |||

RESULT 2
AAS17500
ID AAS17500 standard; cDNA; 390 BP.
XX
XX AAS17500;
XX
XX 14-FEB-2002 (first entry)
XX
XX DNA encoding human beta-like glycoprotein hormone, Beta10.
XX
XX Human: glycoprotein hormone; beta10; homeostatic disorder: diabetes;
KM stress response; immune system dysfunction; tissue damage; cancer;
KM thyroid gland related condition; ss.
XX
XX Homo sapiens.
XX
XX OS
XX
XX Key Location/Qualifiers
FH sig_peptide 1..72
FT /tag- a
FT 1..390
FT CDS
FT /tag- b
FT /product- "beta-like glycoprotein beta10"
FT /partial
FT /note- "No stop codon given"
```

```
FT mat_peptide 73..390
FT /*tag= c
FT /note= "Mature beta10"
XX
XX WO200173034-A2.
XX
XX 04-OCT-2001.
XX
XX 28-MAR-2001; 2001WO-US09999.
XX
XX 28-MAR-2000; 2000US-192654P.
XX 24-APR-2000; 2000US-199211P.
XX 27-NOV-2000; 2000US-0723970.
XX
XX (AMGE-) AMGEN INC.
XX
XX Paszty CJR, Cao J, Danilenko DM, Gong J, Hall DC;
XX WPI; 2002-055150/07.
XX P-PSDB; AAU10366.
XX
XX New polynucleotides encoding polypeptides for treating and diagnosing
XX disorders such as thyroid gland related conditions comprises novel
XX human glycoprotein hormone polypeptide, and the beta10 protein -
XX
XX Claim 1, Fig 1; 201pp; English.
XX
XX The invention relates to an isolated polynucleotide (1) encoding
CC beta-like glycoprotein polypeptide. The polynucleotides, polypeptides and
CC heterodimers can be administered therapeutically (e.g. by administering
CC the polynucleotides to modulate levels of beta10 polypeptide or
CC heterodimer; to treat or prevent diseases and disorders such as
CC homeostatic disorders (e.g. diabetes), disorders related to stress
CC responses (e.g. immune system dysfunction) or disorders requiring
CC increased cell differentiation/proliferation (e.g. tissue damage during
CC cancer treatment). They can be included in compositions which are used
CC for the same purposes. They are useful to diagnose pathological
CC conditions or susceptibility to pathological conditions, especially
CC thyroid gland related conditions. The polypeptides and heterodimers can
CC be used to identify binding compounds. They are useful for producing
CC antibodies, and for detecting or quantifying beta10 polypeptides,
CC heterodimers, and selective binding agents. The polynucleotides can be
CC used to produce cells comprising the polynucleotide, and for producing
CC polypeptides/heterodimers and identifying compounds modulating beta 10
CC polypeptide/heterodimer activity by detecting changes in activity or
CC production in the cell. They can also be used to produce implantation
CC devices to administer polypeptide/heterodimers. The present
CC sequence represents the DNA encoding beta-like glycoprotein hormone
CC beta10.
XX
XX Sequence 390 BP; 69 A; 128 C; 106 G; 87 T; 0 other;
XX

Query Match 100.0%; Score 390; DB 24; Length 390;
Best Local Similarity 100.0%; Pred. No. 4,7e-102;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgaagctggcattctctctctctggcccatgagccctcctctctgctgctatgac 60
    |||
DB 1 atgaagctggcattctctctctctggcccatgagccctcctctctgctgctatgac 60
    |||

QY 61 tgtgtcctcggtgctccagtggaacctgagcacttgggtggtcgtgcgtgaaggag 120
    |||
DB 61 tgtgtcctcggtgctccagtggaacctgagcacttgggtggtcgtgcgtgaaggag 120
    |||

QY 121 ttactctctggccaagaagcagctgcaaggccttcgataccaacagatgctgc 180
    |||
DB 121 ttactctctggccaagaagcagctgcaaggccttcgataccaacagatgctgc 180
    |||

QY 181 tgggtcgtcgtgtagagactgggaaccattctggaacccctatatattgaagccat 240
    |||
DB 181 tgggtcgtcgtgtagagactgggaaccattctggaacccctatatattgaagccat 240
    |||
```

QY 241 catcagagctctacctaacaagagacaaacaggtgactgtcaagctgccccactgtgccc 300
 Db 241 catcagagctctacctaacaagagacaaacaggtgactgtcaagctgccccactgtgccc 300
 QY 301 ccggagagtcgacccctctacactatcccggtgacatccgctgtgtgtggagcctgc 360
 Db 301 ccggagagtcgacccctctacactatcccggtgacatccgctgtgtgtggagcctgc 360
 QY 361 tccactgcacacagagagtgtagacatc 390
 Db 361 tccactgcacacagagagtgtagacatc 390
 RESULT 3
 AAD16347
 ID AAD16347 standard; DNA; 393 BP.
 AC AAD16347;
 XX
 DT 19-NOV-2001 (first entry)
 XX
 DE Human sbhGTA gene.
 XX
 KW Human: Alzheimer's disease; amyotrophic lateral sclerosis;
 KW ALS; Zollinger-Ellison syndrome; immune system disease; schizophrenia;
 KW inflammation; haematopoietic disease; anxiety; feeding disorder; aging;
 KW anorexia; depression; cardiovascular disease; sleep disorder; seizure;
 KW memory alteration; migraine; stroke; asthma; neuropathy; hypoglycaemia;
 KW sexual disorder; growth abnormality; infection; autoimmune disease;
 KW rheumatoid arthritis; cataractogenesis; angiogenesis; atherosclerosis;
 KW cerebral ischaemia; cirrhosis; Huntington's disease; Hodgson's disease;
 KW hypercholesterolaemia; headache; amnesia; cardiac arrhythmia; obesity;
 KW diabetes mellitus; glomerulonephritis; renovascular hypertension;
 KW cancer; vaccine; gene therapy; sbhGTA gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..393
 FT /tag= a
 FT /product= "Human sbhGTA protein"
 XX
 PN W0200160850-A1.
 XX
 PD 23-AUG-2001.
 PD
 PF 14-FEB-2001; 2001MO-US04703.
 PF
 XX 14-FEB-2000; 2000US-0182172.
 PR 29-FEB-2000; 2000US-0186084.
 PR 18-APR-2000; 2000US-0198583.
 PR 04-OCT-2000; 2000US-0237963.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Agarwal P, Kabnick KS, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;
 XX
 DR WPI: 2001-536566/59.
 DR P-PDB: AA09440.
 XX
 PT New secreted and membrane associated polypeptides for treating
 PT Alzheimer's disease, psoriasis, cancer, enterocolitis, sleep and sexual
 PT disorders, stroke, and asthma
 XX
 PS Claim 2; Page 38; 94pp: English.
 XX
 CC The present sequence is a gene encoding human sbhGTA protein,
 CC a secreted protein of the invention.
 CC The invention relates to secreted and membrane associated polypeptides
 CC and nucleic acid molecules encoding such polypeptides. Sequences of the
 CC invention are useful for treating diseases such as Alzheimer's disease,
 CC amyotrophic lateral sclerosis (ALS), Zollinger-Ellison syndrome, diseases

CC of the immune system, haematopoietic disease, inflammation, anxiety,
 CC schizophrenia, feeding disorders, anorexia, depression, social, sexual
 CC and rewarded behaviour, cardiovascular disease, sleep disorder, learning
 CC and memory alteration and altered immune response, seizure, migraine,
 CC cancer, stroke, asthma, neuropathy, aging, sexual disorders, treatment
 CC of transsexuals, growth abnormalities, obesity, infections, autoimmune
 CC diseases (e.g. rheumatoid arthritis), cataractogenesis, angiogenesis,
 CC disorders associated with healthy maintenance of gastric mucosa and
 CC repair of acute and chronic mucosal lesion, lung carcinoma, cerebral
 CC ischaemia, atherosclerosis, cirrhosis, Huntington's disease, headache,
 CC amnesia, multiple sclerosis, Hodgson's disease, dilated cardiomyopathy,
 CC congestive heart failure, cardiac arrhythmias, hypercholesterolaemia,
 CC viral and non-viral hepatitis, type I and type II diabetes mellitus,
 CC glomerulonephritis, renovascular hypertension, hypoglycaemia, periodic
 CC paralyses, tendinitis and malignant hyperthermia. Polypeptides of the
 CC invention are used to identify membrane bound and soluble receptors.
 CC They are also useful as vaccines for inducing an immunological response
 CC in a mammal. Polynucleotides of the invention are used in gene therapy.
 CC They are also valuable for chromosome localisation studies and tissue
 CC expression studies.
 XX
 SO Sequence 393 BP; 70 A; 128 C; 107 G; 88 T; 0 other;
 XX
 Query Match 100.0%; Score 390; DB 22; Length 393;
 Best Local Similarity 100.0%; Pred. No. 4.7e-102;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 atgaagctcgacatctctctctctctgcccacatgacctctctctctgtgtgtgtgc 60
 Db 1 atgaagctcgacatctctctctctctctgcccacatgacctctctctgtgtgtgtgc 60
 QY 61 tgtgcccctgctgctcctcagtgagagctgcgacctgtgtgtgtgtgtgtgtgtgtgtgt 120
 Db 61 tgtgcccctgctgctcctcagtgagagctgcgacctgtgtgtgtgtgtgtgtgtgtgtgt 120
 QY 121 ttacttctctgccaagaagcagctgacagggcctctgcatcaccagagatgcctgc 180
 Db 121 ttacttctctgccaagaagcagctgacagggcctctgcatcaccagagatgcctgc 180
 QY 181 tggggtcgtctgtgagacctgtggaagaaccatctcgaaacccctatatgtgaagccat 240
 Db 181 tggggtcgtctgtgagacctgtggaagaaccatctcgaaacccctatatgtgaagccat 240
 QY 241 catcagagctctacctaacaagagacaaacaggtgactgtcaagctgccccactgtgccc 300
 Db 241 catcagagctctacctaacaagagacaaacaggtgactgtcaagctgccccactgtgccc 300
 QY 301 ccggagagtcgacccctctacactatcccggtgacatccgctgtgtgtggagcctgc 360
 Db 301 ccggagagtcgacccctctacactatcccggtgacatccgctgtgtgtggagcctgc 360
 QY 361 tccactgcacacagagagtgtagacatc 390
 Db 361 tccactgcacacagagagtgtagacatc 390
 RESULT 4
 AAH42567
 ID AAH42567 standard; DNA; 917 BP.
 AC AAH42567;
 XX
 DT 01-OCT-2001 (first entry)
 XX
 DE Nucleotide sequence of a human cystine knot polypeptide.
 XX
 KW Cystine knot polypeptide; follicular arrest; recruitment modulator;
 KW fertility-related disorder; contraception; menopause; contraceptive;
 KW follicle growth, ss.
 XX
 OS Homo sapiens.
 XX

Key Location/Qualifiers
CDS 101..493
/*tag= a
/product= "cystine knot polypeptide"

MO200153346-A1.

26-JUL-2001.

17-JAN-2001; 2001WO-EP00570.

18-JAN-2000; 2000EP-0200185.

(ALKD) AKZO NOBEL NV.

Moselman S, Spek Van Der Pj;

WPI: 2001-476102/51.

P-PSDB: AAG63211.

New DNA sequences, useful for coding or producing cystine knot polypeptides, which are useful in preparing a pharmaceutical for fertility-related disorders or contraception, and for controlling follicular arrest and recruitment.

Claim 6; Page 23; 29pp; English.

The present sequence encodes a human cystine knot polypeptide. The polypeptide is a follicular arrest and recruitment modulator. Cystine knot polypeptides are useful in preparing a pharmaceutical for fertility-related disorders or in contraception. The polypeptide is particularly useful for controlling follicular arrest and recruitment. Inhibition of recruitment can be used to delay (premature) menopause or as a contraceptive. The polypeptide is also useful for in vitro maturation and growth of follicles, e.g. from frozen ovarian tissue.

Sequence 917 BP; 222 A; 249 C; 205 G; 241 T; 0 other:

Query Match 100.0%; Score 390; DB 22; Length 917;
Best Local Similarity 100.0%; Pred. No. 6.1e-102;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgaagctgagcattcctctcctgagcccatgagccctcctcctctctgctgctatgac 60
|||||
DB 101 atgaagctgagcattcctctcctgagcccatgagccctcctcctctctgctgctatgac 160
|||||

QY 61 tgtgtccctcgtgctcctcagtgaggagccttgagccttctgtgagctgtgagggag 120
|||||
DB 161 tgtgtccctcgtgctcctcagtgaggagccttgagccttctgtgagctgtgagggag 220
|||||

QY 121 ttattcttcctgagcaagaagctgagcagggccttcctgagatcacacagagatgctgc 180
|||||
DB 221 ttattcttcctgagcaagaagctgagcagggccttcctgagatcacacagagatgctgc 280
|||||

QY 181 tggggctcgttgagaccttgaggagaaccattctggaacccctcctatataatgaagccat 240
|||||
DB 281 tggggctcgttgagaccttgaggagaaccattctggaacccctcctatataatgaagccat 340
|||||

QY 241 catcagagctgagctacacagacgaagagtgagctgagctgaagctgcccagctgccc 300
|||||
DB 341 catcagagctgagctacacagacgaagagtgagctgagctgaagctgcccagctgccc 400
|||||

QY 301 ccggagagctgagacccctctacacataccctgtgagcattcctgctgagctgagagcctgc 360
|||||
DB 401 ccggagagctgagacccctctacacataccctgtgagcattcctgctgagctgagagcctgc 460
|||||

QY 361 tccactgacacacagagagtgtagacatc 390
|||||
DB 461 tccactgacacacagagagtgtagacatc 490
|||||

AAf83867
ID AAF83867 standard; cDNA: 693 BP.
XX
AC AAF83867;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human novel gonadotropin (NOVGON) protein encoding cDNA.
XX
KW NOVG: transmembrane protein; NOVTRAN; neuromedin peptide; NOVNEUR;
gonadotropin-like protein; NOVGON; interleukin-1; NOVINTRA; human;
cytostatic; neuroprotective; reproductive; antiinflammatory; cancer;
antibacterial; cerebroprotective; antidiabetic; antiarthritic;
antiallergic; antiallergic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..693
CDS /*tag= a
/product= "NOVGON"

MO200140291-A2.

07-JUN-2001.

06-DEC-2000; 2000WO-US33029.

06-DEC-1999; 99US-0169056.
09-DEC-1999; 99US-0169866.
09-DEC-1999; 99US-0169886.
10-DEC-1999; 99US-0170252.
12-JAN-2000; 2000US-0175740.
05-DEC-2000; 2000US-0170252.

(CURA-) CURAGEN CORP.

Burgess CE, Prayaga SK, Shinkels RA, Rastelli L, Zernhusen BD;
Mezes PS;
WPI: 2001-374790/39.
P-PSDB: AAB84998.

Novel isolated human transmembrane, neuromedin peptide
gonadotropin-like protein and interleukin-1 receptor antagonist
proteins, useful for treating cancer, immune response disorder,
metabolic function disorders

Claim 8; Fig 6A; 138pp; English.

The invention provides novel polypeptides (NOV) selected from human
transmembrane protein (NOVTRAN), neuromedin peptide (NOVNEUR),
gonadotropin-like protein (NOVGON) and two interleukin-1 receptor
antagonist proteins (NOVINTRA A and B). The invention also provides
methods in which a NOV polypeptide, polynucleotide and antibody are
used in the detection, prevention and treatment of a broad range of
pathological states. NOVTRAN can be used to treat a cell signaling
disorder such as cancer, immune response disorder, hematopoietic
disorder, neurodegenerative disorder. NOVNEUR can be used to treat
endocrine disorder, muscle disorder, neurologic disorder, cancers of
central nervous system, breast, colon, ovary, kidney, prostate and
thyroid. NOVGON can be used to treat reproductive development disorder,
metabolic function disorder and melanoma. NOVINTRA A and B can be used
to treat bone metabolism or structure disorder, inflammatory response
disorder, immune regulation disorder, septic shock, stroke, diabetes,
arthritis and cancer. The present sequence represents a cDNA encoding
the NOVGON polypeptide.

Sequence 693 BP; 138 A; 197 C; 196 G; 162 T; 0 other;

Query Match 96.9%; Score 378; DB 22; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.5e-96;

Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgaagctggcatcctctctcttgcccataggccctctctctctctgctgctatgac 60
Db 1 atgaagctggcatcctctctcttgcccataggccctctctctctgctgctatgac 60

QY 61 tgtctctggtgctcctcagtggaacctgcgaaccttgtggctgtgccgtgaaggag 120
Db 61 tgtctctggtgctcctcagtggaacctgcgaaccttgtggctgtgccgtgaaggag 120

QY 121 tttaactctctggccaagaagcaggtctgaaggagccttgatcaccaagatgctgc 180
Db 121 tttaactctctggccaagaagcaggtctgaaggagccttgatcaccaagatgctgc 180

QY 181 tggagctgctgtgagacctggagaaccatctctggaacccccctatatgtgaagccat 240
Db 181 tggagctgctgtgagacctggagaaccatctctggaacccccctatatgtgaagccat 240

QY 241 catcgagctgtacctacaacgagacaaacaggtgactgtcaagctgcccactgtgc 300
Db 241 catcgagctgtacctacaacgagacaaacaggtgactgtcaagctgcccactgtgc 300

QY 301 ccgggaagtcgagccctctacactatccgtggcactcgcgtgtgactgctggagcctgc 360
Db 301 ccgggaagtcgagccctctacactatccgtggcactcgcgtgtgactgctggagcctgc 360

QY 361 tccactgccaccacgagag 378
Db 361 tccactgccaccacgagag 378

RESULT 6
AAH46589
ID AAH46589 standard; DNA: 318 BP.
XX
AC AAH46589;
XX
DT 17-SEP-2001 (first entry)
XX
DE Human anterior pituitary hormone-related polynucleotide #2.
XX
KW Human; anterior pituitary hormone; hypertension; autoimmune disease;
KW heart failure; ds.
XX
OS Homo sapiens.
XX
PN WO200144475-A1.
PD 21-JUN-2001.
XX
PF 15-DEC-2000; 2000MO-JP08896.
XX
PR 17-DEC-1999; 99JP-0358707.
PR 18-FEB-2000; 2000JP-0046825.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Fukusumi S, Fujii R, Hosoya M;
XX
DR WPI; 2001-408485/43.
DR P-PSDB; AAG64065.
XX
PT Polypeptides for treatment of hypertension, autoimmune disease and
XX heart failure -
XX
PS Claim 7; Page 100-101; 107pp; Japanese.
XX
CC The invention relates to a novel polypeptide comprising a fully defined
CC 130 amino acid sequence given in the specification and its amides,
CC esters and salts. The polypeptide has anterior pituitary hormone-related
CC activity. It is useful for the treatment of hypertension, autoimmune
CC diseases and heart failure. The screening method and kit also
CC provided in the invention are useful for identifying new substances

CC for treating and preventing these diseases. The present sequence
CC encodes a polypeptide of the invention.
XX
SQ Sequence 318 BP; 63 A; 103 C; 88 G; 64 T; 0 other;

Query Match 81.5%; Score 318; DB 22; Length 318;
Best Local Similarity 100.0%; Pred. No. 1.8e-81;
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 gccctcagttggaaacctgcgcacaccttggctgtgctgtgcgtgagggagttacttctg 132
Db 1 gccctcagttggaaacctgcgcacaccttggctgtgctgtgcgtgagggagttacttctg 60

QY 133 gccaaagaagcgaagctgcaggggacctcgatcaccaacagatgcctgtgggtgcgtg 192
Db 61 gccaaagaagcgaagctgcaggggacctcgatcaccaacagatgcctgtgggtgcgtg 120

QY 193 gagaacctggagaaccacattctggaacccccctatatgtgaagccatcatcgagtcgt 252
Db 121 gagaacctggagaaccacattctggaacccccctatatgtgaagccatcatcgagtcgt 180

QY 253 acctacaacgagacccaacaggtgactgtcaagctgtgccaaactgtgcccggagtcgac 312
Db 181 acctacaacgagacccaacaggtgactgtcaagctgtgccaaactgtgcccggagtcgac 240

QY 313 cccttctaacactatccgctggccatccgctgtgactgtgagctgtcactgtccacc 372
Db 241 cccttctaacactatccgctggccatccgctgtgactgtgagctgtcactgtccacc 300

QY 373 acggaagtgtgaaccalc 390
Db 301 acggaagtgtgaaccalc 318

RESULT 7
AAS17508
ID AAS17508 standard; cDNA: 393 BP.
XX
AC AAS17508;
XX
DT 14-FEB-2002 (first entry)
XX
DE DNA encoding mouse beta-like glycoprotein hormone, Beta10.
XX
KW Mouse; glycoprotein hormone; beta10; homeostatic disorder; diabetes;
KW stress response; immune system dysfunction; tissue damage; cancer;
KW thyroid gland related condition; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..69 /*tag- a
FT CDS 1..393 /*tag- b
FT mat_peptide /product= "beta-like glycoprotein beta10"
FT 70..390 /*tag- c
FT /note= "Mature beta10"
XX
XX
XX WO200173034-A2.
XX
XX 04-OCT-2001.
XX
XX 28-MAR-2001; 2001MO-US09999.
XX
XX 28-MAR-2000; 2000US-192654P.
XX 24-APR-2000; 2000US-199211P.
XX 27-NOV-2000; 2000US-0723970.
XX
XX (AMGE-) AMGEN INC.
XX

QY 361 tcaactgcacacagaggtgtgagaccatc 390
 Db 358 tccactgcccacactgagtgtagaccatc 387

RESULT 9
 ID AAH46594 standard; DNA; 318 BP.
 AC AAH46594;
 XX

DT 17-SEP-2001 (first entry)

DE Rat anterior pituitary hormone-related polynucleotide #2.

KM Rat; anterior pituitary hormone; hypertension; autoimmune disease;
 KW heart failure; ds.

OS Rattus sp.

XX WO200144475-A1.

PD 21-JUN-2001.

PF 15-DEC-2000; 2000WO-JP08896.

XX 17-DEC-1999; 99JP-0358707.

PR 18-FEB-2000; 2000JP-0046825.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hinuma S, Fukusumi S, Fujii R, Hosoya M;

PI WPI; 2001-408485/43.

DR P-PSDB; AAG64068.

XX Polypeptides for treatment of hypertension, autoimmune disease and

PT heart failure -

PS Claim 7; Page 104; 107pp; Japanese.

CC The invention relates to a novel polypeptide comprising a fully defined
 CC 130 amino acid sequence given in the specification and its amides,
 CC esters and salts. The polypeptide has anterior pituitary hormone-related
 CC activity. It is useful for the treatment of hypertension, autoimmune
 CC diseases and heart failure. The screening method and kit also
 CC provided in the invention are useful for identifying new substances
 CC for treating and preventing these diseases. The present sequence is
 CC is a polynucleotide encoding a polypeptide provided in the specification.

XX Sequence 318 BP; 69 A; 89 C; 90 G; 70 T; 0 other;

Query Match 66.7%; Score 260; DB 22; Length 318;

Best Local Similarity 88.9%; Pred. No. 7.2e-65; Mismatches 35; Indels 0; Gaps 0;

Matches 281; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 75 ctccagtgaggaaacctgtgcacctgtgtggctgtccgtgagggagttacttctcggc 134

Db 3 ctccagtgaggaaacctgtgtggctgtgtggctgtgtggctgtgtggctgtgtggc 62

QY 135 caagaagcagagctgtcagggccttgatcaccaagagatgctgtcgtgggtcgttga 194

Db 63 caagaagcagagctgtcagggccttgatcaccaagagatgctgtcgtgggtcgttga 122

QY 195 gactctggagaagaacctcttggaacccctatatattgaagccatcatcagagctgtac 254

Db 123 gactctggagaagaacctcttggaacccctatatattgaagccatcatcagagctgtac 182

QY 255 ctacaagagacacaaagagtgactgtcaagctgtcccaactgtgtcccgaggagtcgagccc 314

Db 183 ctacaagagagacacaaagagtgactgtcaagctgtcccaactgtgtcccgaggagtcgagccc 242

QY 315 ctctacacatattccgtgacatccgtctgtgactgcgagagctgtctcaactgcaccac 374
 Db 243 ctctacacatattccgtgacatccgtctgtgactgcgagagctgtctcaactgcaccac 302

QY 375 ggagtgtagaacatc 390
 Db 303 tgagtgtagaacatc 318

RESULT 10

AAH42568
 ID AAH42568 standard; DNA; 1045 BP.

AC AAH42568;
 XX

DT 01-OCT-2001 (first entry)

DE Nucleotide sequence of human cystine knot polypeptide splice variant.

XX Cystine knot polypeptide; follicular arrest; recruitment modulator;
 KW fertility-related disorder; contraception; menopause; contraceptive;
 KM follicle growth; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

FT CDS 101..328

FT /*tag= a

FT /*product= "cystine knot polypeptide"

XX WO200153346-A1.

PD 26-JUL-2001.

PF 17-JAN-2001; 2001WO-EP00570.

XX 18-JAN-2000; 2000EP-0200185.

XX (ALKU) AKZO NOBEL NV.

PI Mosselman S, Spek Van Der PJ;

XX WPI; 2001-476102/51.

DR P-PSDB; AAG63212.

XX New DNA sequences, useful for coding or producing cystine knot

XX polypeptides, which are useful in preparing a pharmaceutical for

XX follicular arrest and recruitment -

PS Claim 7; Page 24; 29pp; English.

XX The present sequence encodes a human cystine knot polypeptide splice

XX variant. The polypeptide is a follicular arrest and recruitment

XX modulator. Cystine knot polypeptides are useful in preparing a

XX pharmaceutical for fertility-related disorders or in contraception.

XX The polypeptide is particularly useful for controlling follicular

XX arrest and recruitment. Inhibition of recruitment can be used to delay

XX (premature) menopause or as a contraceptive. The polypeptide is also

XX useful for in vitro maturation and growth of follicles, e.g. from

XX frozen ovarian tissue.

Query Match 64.6%; Score 252; DB 22; Length 1045;

Best Local Similarity 75.3%; Pred. No. 2e-62; Mismatches 128; Gaps 1;

Matches 390; Conservative 0; Mismatches 128; Indels 128; Gaps 1;

QY 1 atgaagtgcatctctctctctgtgagccatgagccctctctctctctctctctgtgtatgac 60

Db 101 atgaagtgcatctctctctctgtgagccatgagccctctctctctctctctgtgtatgac 160

```
OY 61 tttgtctctggtgctccagctgggaacctgcgcacctttgttggtctgtccgtgaggag 120
    |||
Db 161 tctgtctctggtgctccagctgggaacctgcgcacctttgttggtctgtccgtgaggag 220
OY 121 ttacttctctggccaagaagcagagctgcaagggtcttcgagtcacacacagatgctcgc 180
    |||
Db 221 ttacttctctggccaagaagcagagctgcaagggtcttcgagtcacacacagatgctcgc 280
OY 181 tgggtgctgctgtgagaacctggagag----- 204
    |||
Db 281 tgggtgctgctgtgagaacctggagagctttgtcaagatgtctgtatgaaagaagcatca 340
OY 205 ----- 204
Db 341 atacacattgttgttgtagctgagatgcctccctcgtgagctgtagatctccagcct 400
OY 205 -----aaaccattctggaacccctatatgt 232
    |||
Db 401 aatggaagcgcatttagaatcacactgtgcactaaaccattcggaaacccctatatgt 460
OY 233 aagcccatcatcgagctctgtacactacaagagaacaaagaagtgcactgtcaagctgccca 292
    |||
Db 461 aagcccatcatcgagctctgtacactacaagagaacaaagaagtgcactgtcaagctgccca 520
OY 293 actgtgccccggagctgcagacccttctacacatcccggtggcaccgtgtgactgcg 352
    |||
Db 521 actgtgccccggagctgcagacccttctacacatcccggtggcaccgtgtgactgcg 580
OY 353 gagcctgctcactgcgcacacagagagtgtagaacatc 390
    |||
Db 581 gagcctgctcactgcgcacacagagagtgtagaacatc 618

RESULT 11
AAS17518
ID AAS17518 standard; DNA: 2985 BP.
XX
AC AAS17518;
XX
DT 14-FEB-2002 (first entry)
XX
DE Genomic DNA encoding mouse beta-1-like glycoprotein hormone, beta10.
XX
KM Mouse; glycoprotein hormone; beta10; homeostatic disorder: diabetes;
KW stress response; immune system dysfunction; tissue damage; cancer;
XX thyroid gland related condition; ds.
OS
XX Mus musculus.
XX
PN WO200173034-A2.
XX
PD 04-OCT-2001.
XX
PF 28-MAR-2001; 2001WO-US09999.
XX
PR 28-MAR-2000; 2000US-192654P.
PR 24-APR-2000; 2000US-199211P.
PR 27-NOV-2000; 2000US-0723970.
XX
PA (AMGE-) AMGEN INC.
XX
PI Paszty CJR, Cao J, Danilenko DM, Gong J, Hill DC;
XX
DR WPI; 2002-055150/07.
XX
PT New polynucleotides encoding polypeptides for treating and diagnosing
PT disorders such as thyroid gland related conditions comprises novel
PT human glycoprotein hormone polypeptide, and the beta10 protein -
XX
XX Example 6; Page 198-199; 201pp; English.
XX
CC The invention relates to an isolated polynucleotide (I) encoding
CC beta-1-like glycoprotein polypeptide. The polynucleotides, polypeptides and
```

```
CC heterodimers can be administered therapeutically (e.g. by administering
CC the polynucleotides to modulate levels of beta10 polypeptide or
CC heterodimer; to treat or prevent diseases and disorders such as
CC homeostatic disorders (e.g. diabetes), disorders related to stress
CC responses (e.g. immune system dysfunction) or disorders requiring
CC increased cell differentiation/proliferation (e.g. tissue damage during
CC cancer treatment). They can be included in compositions which are used
CC for the same purposes. They are useful to diagnose pathological
CC conditions or susceptibility to pathological conditions, especially
CC thyroid gland related conditions. The polypeptides and heterodimers can
CC be used to identify binding compounds. They are useful for producing
CC antibodies, and for detecting or quantifying beta10 polypeptides,
CC heterodimers, and selective binding agents. The polynucleotides can be
CC used to produce cells comprising the polynucleotide, and for producing
CC polypeptides/heterodimers and identifying compounds modulating beta 10
CC polypeptide/heterodimer activity by detecting changes in activity or
CC production in the cell. They can also be used to produce implantation
CC devices to administer polypeptide/heterodimers. The present sequence
CC represents the genomic DNA encoding mouse beta-1-like glycoprotein hormone
CC beta10.
XX
SQ Sequence 2985 BP; 784 A; 630 C; 761 G; 810 T; 0 other;
```

Query Match 40.0%; Score 156; DB 24; Length 2985;

Best Local Similarity 85.3%; Pred. No. 8, 4e-35; Matches 174; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```
OY 1 atgaagctgcaattctctctcttctgcccacatggccctctctctctgtgctgatagc 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 atgaagctgatacttctctcttctgtgcagtgccctctctctctctgtggtgcctgac 60
OY 61 tgtgtctcgtgtgctctcagctgggaacctgcgacctgtgtggctgtgcgtgaggag 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 tctgtcctcagcagctccagctgggaacctgcacaccttgtgtggctgtgtgaggaga 120
OY 121 ttacttctctggccaagaagcagagctgcagggccttcgatacacacagatgctgc 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 ttacttctatggccaagaagcagagctgcagggccttcgatacacacagatgctgc 180
OY 181 tgggtgctgctgtgagacctggag 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 tgggtgctgctgtgagacctggag 204
```

```
RESULT 12
AAN60523
ID AAN60523 standard; cDNA: 496 BP.
XX
AC AAN60523;
XX
DT 01-JUL-1991 (first entry)
XX
DE Sequence encoding dog beta luteinising hormone (LH).
XX
KM Contraceptive; fertility control; vaccine; reproductive hormone; ss.
XX
OS Dog.
XX
FH Key 1.54 Location/Qualifiers
FH CDS 55..420 -
FT mat_peptide /*tag= a
FT polyA_signal /*tag= b
FT polyA_site /*tag= c
FT /*tag= d
XX
PN WO8607383-A.
XX
PD 18-DEC-1986.
XX
```

PF 04-JUN-1986; 86WO-US01226.
XX 18-JUL-1985; 85US-0756847.
PR 04-JUN-1985; 85US-0741168.
XX
PA (BIOT-) BIOTECHN RES PARTN.
XX
PI Talmadge KD, Fildes JC;
XX
DR WPI; 1986-346608/52.
PR P-PSDB; AAP60601.
XX
PT Auto-antigen vaccines conferring antigenicity using multimers
XX etc. - useful as species specific or cross-species effective,
XX esp. for controlling fertility in mammals
PS Example; Fig 3B; 101pp; English.
XX
CC The patentors claim a vaccine effective against mammalian fertility
CC comprising a vaccinia virus genome having disposed unit, in a non-
CC essential region, a DNA sequence of formula : (Hormone)_n; n=1-20;
CC Hormone= DNA sequence derived from the sequence encoding a
CC reproductive hormone. The hormone is esp. LH, GnRH, CG or FSH.
XX
SQ Sequence 496 BP; 66 A; 191 C; 134 G; 105 T; 0 other;

Query Match 15.8%; Score 61.6; DB 7; Length 496;
Best Local Similarity 52.8%; Pred. No. 5.1e-08;
Matches 133; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 133 gccaaagacgagctgtagggccctcgatcaacacgagatgctgtggtcgtgt 192
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 106 gctgagacgagcgtgctgtgtatcaacttcacacacacatctgtgcggtac 165
QY 193 gagacctgggagaacccatcttggaacccccctatatgagccatcagatctgt 252
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 166 tgcacacagatgtagtctgtccagcgcgcctgacactgtgtccacacacatgtgc 225
QY 253 acctacacgagacacacagtgtagctgcaagctgcccactgtgcccggagatcgac 312
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 226 acctacacatgtagctgtagctgtgtatcaatccggtcccgatgcccgtggtgac 285
QY 313 cccctctacacatcccgctggtgacatccgctgtgtagctggagacgtgtccactgacac 372
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 286 cccaatgtctcctccctccgctggtccctcagctgtgctgtgggcccgtgtcagcaac 345
QY 373 acgagatgtgag 384
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 346 tccgactgtggg 357

RESULT 13
AAQ03844
ID AAQ03844 standard; cDNA; 525 BP.
XX
AC AAQ03844;
XX
DT 24-AUG-1990 (first entry)
XX
DE Bovine beta LH subunit.
XX
KW Luteinizing hormone; follicle stimulating hormone;
KW recombinant cDNA; alpha subunit; beta subunit; ungulate; ss.
XX
OS Bos taurus.
XX
FH Key
FT CDS
FT
FT
XX
PN WO9002757-A.

XX 22-MAR-1990.
PD
XX 02-SEP-1988; 88WO-0030949.
PF
XX 02-SEP-1988; 88WO-US03049.
PR
XX (INTE-) INTEGRATED GENETICS.
XX
PI Beck A, Bernstein E, Hsiung N, Kelton C, Lerner T, Reddy VB;
PI Chapel SC;
XX
DR WPI; 1990-115954/15.
PR
XX Biologically active ungulate LH and FSH- produced by recombinant methods.
XX
PS Disclosure; Fig 2; 66pp; English.
XX
XX LH and FSH comprises an alpha and a beta subunit, both subunits can be
CC synthesised in a single cell conty. an expression vector comprising
CC heterologous DNA encoding one subunit.
CC See also AAQ03843-Q03851.
XX
SQ Sequence 525 BP; 90 A; 202 C; 130 G; 103 T; 0 other;

Query Match 14.8%; Score 57.8; DB 11; Length 525;
Best Local Similarity 51.8%; Pred. No. 6.3e-07;
Matches 131; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 132 gccaaagacgagctgtagggccctcgatcaacacgagatgctgtggtcgtgt 191
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 107 gctgagagagagcgtgctgtgtatcaacttcacacacacatctgtgcggtac 166
QY 192 tgaacctgtggagaacccatcttggaacccccctatatgagccatcagatctgt 251
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 167 ctgcacacagatgagacggtgtgctgtatcctgtccgcacatgcccacgaggtgtg 226
QY 252 tactacacgagacacacacagtgtagctgcaagctgcccactgtgcccggagatcgac 311
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 227 cactacacatgtagctgtgctgtgtcgtccctcggtcgtcccggtgcccacatgtagtga 286
QY 312 cccctctacacatcccgctggtgacatccgctgtgtagctggagacgtgtccactgacac 371
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 287 cccaatgtctcctccctccgctggtccctcagctgtcactgtgtagctgtgacccgtccctcagca 346
QY 372 cagagatgtgag 384
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 347 cactgactgcggg 359

RESULT 14
AAH41044
ID AAH41044 standard; DNA; 535 BP.
XX
AC AAH41044;
XX
DT 28-AUG-2001 (first entry)
XX
DE GTH-beta LH subunit DNA.
XX
KW Polymeric glycoprotein; goldfish; GTH; gonadotrophin; ds.
KW Carassius auratus.
XX
OS
XX JP2001086992-A.
XX
PD 03-APR-2001.
XX
PF 24-SEP-1999; 99JP-0270790.
XX
PR 24-SEP-1999; 99JP-0270790.
XX
XX

Location/Qualifiers
3..422
/*tag= a
/product=Bovine beta LH

PA (AJIN) AJINOMOTO KK.
XX
XX WPI, 2001-338422/36.
DR
XX
XX Production of a polymeric glycoprotein by introducing a gene expressing
PT the glycoprotein into a fish -
XX
XX
PS Example 1; Page 9-10; 14pp; Japanese.
XX
XX The present invention relates to a method for the production of a
CC polymeric glycoprotein in which a gene encoding the glycoprotein is
CC introduced into a fish and expressed. The glycoprotein can be recovered
CC from the embryo of the fish in which the gene was introduced or from its
CC tissue or blood. The method is used for the production of a polymeric
CC glycoprotein. The present sequence represents the goldfish GTH
CC (gonadotropin) beta1 subunit. DNA encoding GTH is used in an example
XX illustrating the method of the invention.
SO Sequence 535 BP; 135 A; 141 C; 113 G; 146 T; 0 other;

Query Match 14.6%; Score 56.8; DB 22; Length 535;
Best Local Similarity 59.1%; Pred. No. 1.2e-06;
Matches 97; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 222 cccctatatgaagccatcatcagctctgtaactaacacgagacaaacagtgactgt 281
DB 223 cccatttcacactgtccacacacatgtgtcactaccggagcgtgcacgagactgt 282
QY 282 caagctgcccacactgtcccgagagtcgacccctctacacatcccgctggccatcg 341
DB 283 ccgcttcacagactgtctctcccaagggtgagacccacacatacctactcctgtgtcctcag 342
QY 342 ctgtgactgcgagcctgtctcactgcacacacgagagtgatga 385
DB 343 ctgcgactgcagcctgtgcactatgacacatcgtactga 386

RESULT 15
AAI6195
ID AAI6195 standard; DNA; 5651 BP.
XX
XX AAI6195;
AC
XX
XX 29-JAN-2002 (first entry)
DT
XX
XX Human FSH beta encoding plasmid pBYDH1022 SEQ ID NO 6.
DE
XX
XX Human: FSH alpha; FSH beta; follicle stimulating hormone; glycosylation;
KW antiinfertility; cyclic; circular; ds.
KM
XX
XX Chimeric - Homo sapiens.
OS
OS Chimeric - Cytomegalovirus.
OS
XX
XX Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1231..1620
FT /*tag= a
FT /product= "FSH beta"
FT sig_peptide 1231..1284
FT /*tag= b
FT mat_peptide 1285..1617
FT /*tag= c
FT /product= "FSH beta"
XX
XX
XX WO200158493-A1.
XX
XX 16-AUG-2001.
PD
XX
XX 09-FEB-2001; 2001WO-DK00090.
PF
XX
XX 11-FEB-2000; 2000DK-0000220.
PR
XX
XX 14-JUL-2000; 2000DK-0001092.
PR

XX
PA (MAXY-) MAXYGEN APS.
XX
XX Schambye HT, Andersen KV, Van Den Hazel B, Christiansen J;
PI Jeppesen CB;
XX
XX WPI, 2001-607186/59.
DR
XX
XX
XX New polypeptide conjugate with follicle stimulating hormone (FSH)
PT activity, used to treat infertility, comprises polypeptide having
PT modified FSH alpha and beta subunits with attachment group for
PT non-polypeptide moiety -
XX
XX
PS Example 1; Page 77-79; 88pp; English.
XX
XX The invention relates to new polypeptides and polypeptide conjugates
CC exhibiting follicle stimulating hormone (FSH) activity. FSH is a dimeric
CC hormone comprising of an alpha (AAM51709) and beta (AAM51711) subunit.
CC The invention relates to a heterodimeric FSH conjugate comprising, a
CC dimeric polypeptide having FSH alpha (AAM51733-AAM51800,
CC AAM51442-AAM51449) and beta (AAM52001-AAM52104) subunits, where at least
CC one subunit differs from corresponding wild-type subunit in that an
CC amino acid residue containing an attachment group for a non-polypeptide
CC molecule has been introduced or removed, especially where at least one of
CC the FSH-alpha and FSH-beta subunits comprises at least one introduced
CC N- or O-glycosylation site at its N-terminal and the glycosylation site
CC being glycosylated. The polypeptides have antiinfertility activity. The
CC polypeptides have increased functional in vivo half life and/or serum
CC half life as compared to human FSH, replenishing insufficient endogenous
CC FSH production in a patient. The present sequence is that of a plasmid
CC for expression of human FSH beta.
CC Note: The present sequence differs from that given as the full length FSH
CC beta in SEQ ID NO 3 (AAM51711) as there is a Lys to Glu mutation at
CC position 2 of the native signal sequence.
XX
SO Sequence 5651 BP; 1315 A; 1493 C; 1439 G; 1404 T; 0 other;

Query Match 14.6%; Score 56.8; DB 22; Length 5651;
Best Local Similarity 64.4%; Pred. No. 2.5e-06;
Matches 85; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 249 ctgtactacaagagacccaacagtgactgtcaagctgtcccaactgtgccccggagat 308
DB 1434 ctgcacctcaagagactgtgtacgagacgctgcgggtgccggctgcgccacacgc 1493
QY 309 cgaccctctacacatccgctggccatccgctgtgactgcgagcgccgtccactgc 368
DB 1494 cgacagctgttacacacaccctccgctggccacccactgtccactgtgcgaagtgcgaacgcga 1553
QY 369 caccacgagatg 380
DB 1554 cagcagcagctg 1565

Search completed: July 8, 2002, 12:57:51
Job time: 56263 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2002, 21:20:37 ; Search time 58.55 Seconds
(without alignments)
1636.158 Million cell updates/sec

Title: US-09-723-970-2

Perfect score: 390
Sequence: 1 atgaagctggcattcctctt.....ccacgagtgtagaccatc 390

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1na/5a.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5b.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6a.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6b.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------------|-------------------|
| 1 | 48.8 | 12.5 | 681 | 4 US-08-918-288-31 | Sequence 31, Appl |
| 2 | 48.8 | 12.5 | 681 | 4 US-09-282-357-31 | Sequence 31, Appl |
| 3 | 48.8 | 12.5 | 693 | 4 US-08-918-288-28 | Sequence 28, Appl |
| 4 | 48.8 | 12.5 | 693 | 4 US-09-282-357-28 | Sequence 28, Appl |
| 5 | 48.8 | 12.5 | 702 | 4 US-08-918-288-13 | Sequence 13, Appl |
| 6 | 48.8 | 12.5 | 702 | 4 US-09-282-357-13 | Sequence 13, Appl |
| 7 | 48.8 | 12.5 | 707 | 4 US-08-918-288-29 | Sequence 29, Appl |
| 8 | 48.8 | 12.5 | 707 | 4 US-09-282-357-29 | Sequence 29, Appl |
| 9 | 48.8 | 12.5 | 719 | 4 US-08-918-288-26 | Sequence 26, Appl |
| 10 | 48.8 | 12.5 | 719 | 4 US-09-282-357-26 | Sequence 26, Appl |
| 11 | 48.8 | 12.5 | 728 | 4 US-08-918-288-11 | Sequence 11, Appl |
| 12 | 48.8 | 12.5 | 728 | 4 US-09-282-357-11 | Sequence 11, Appl |
| 13 | 48.8 | 12.5 | 1006 | 1 US-08-239-256-1 | Sequence 1, Appl |
| 14 | 46.4 | 11.9 | 718 | 4 US-08-918-288-10 | Sequence 10, Appl |
| 15 | 46.4 | 11.9 | 718 | 4 US-09-282-357-10 | Sequence 10, Appl |
| 16 | 46.4 | 11.9 | 744 | 4 US-08-918-288-8 | Sequence 8, Appl |
| 17 | 46.4 | 11.9 | 744 | 4 US-09-282-357-8 | Sequence 8, Appl |
| 18 | 45 | 11.5 | 726 | 4 US-08-918-288-16 | Sequence 16, Appl |
| 19 | 45 | 11.5 | 726 | 4 US-09-282-357-16 | Sequence 16, Appl |
| 20 | 45 | 11.5 | 752 | 4 US-08-918-288-14 | Sequence 14, Appl |
| 21 | 45 | 11.5 | 752 | 4 US-09-282-357-14 | Sequence 14, Appl |
| 22 | 43.2 | 11.1 | 549 | 4 US-08-918-288-37 | Sequence 37, Appl |
| 23 | 43.2 | 11.1 | 549 | 4 US-09-282-357-37 | Sequence 37, Appl |
| 24 | 43.2 | 11.1 | 575 | 4 US-08-918-288-35 | Sequence 35, Appl |
| 25 | 43.2 | 11.1 | 575 | 4 US-09-282-357-35 | Sequence 35, Appl |
| 26 | 43.2 | 11.1 | 717 | 4 US-08-918-288-7 | Sequence 7, Appl |
| 27 | 43.2 | 11.1 | 717 | 4 US-09-282-357-7 | Sequence 7, Appl |

| | | | | | |
|----|------|------|------|--------------------|-------------------|
| 28 | 43.2 | 11.1 | 743 | 4 US-08-918-288-5 | Sequence 5, Appl |
| 29 | 43.2 | 11.1 | 743 | 4 US-09-282-357-5 | Sequence 5, Appl |
| 30 | 43.2 | 11.1 | 834 | 4 US-08-918-288-4 | Sequence 4, Appl |
| 31 | 43.2 | 11.1 | 834 | 4 US-09-282-357-4 | Sequence 4, Appl |
| 32 | 43.2 | 11.1 | 835 | 4 US-08-918-288-40 | Sequence 40, Appl |
| 33 | 43.2 | 11.1 | 835 | 4 US-09-282-357-40 | Sequence 40, Appl |
| 34 | 43.2 | 11.1 | 836 | 4 US-08-918-288-2 | Sequence 2, Appl |
| 35 | 43.2 | 11.1 | 836 | 4 US-09-282-357-2 | Sequence 2, Appl |
| 36 | 43.2 | 11.1 | 837 | 4 US-08-918-288-38 | Sequence 38, Appl |
| 37 | 43.2 | 11.1 | 837 | 4 US-09-282-357-38 | Sequence 38, Appl |
| 38 | 43.2 | 11.1 | 893 | 4 US-09-142-320-3 | Sequence 3, Appl |
| 39 | 43.2 | 11.1 | 893 | 4 US-09-142-320-5 | Sequence 5, Appl |
| 40 | 43.2 | 11.1 | 893 | 4 US-09-142-320-6 | Sequence 6, Appl |
| 41 | 43.2 | 11.1 | 893 | 4 US-09-142-320-7 | Sequence 7, Appl |
| 42 | 43.2 | 11.1 | 893 | 4 US-09-142-320-8 | Sequence 8, Appl |
| 43 | 43.2 | 11.1 | 893 | 4 US-09-142-320-9 | Sequence 9, Appl |
| 44 | 43.2 | 11.1 | 893 | 4 US-09-142-320-10 | Sequence 10, Appl |
| 45 | 43.2 | 11.1 | 1301 | 4 US-08-804-166-7 | Sequence 7, Appl |

ALIGNMENTS

RESULT 1
US-08-918-288-31/c
Sequence 31, Application US/08918288
Patent No. 6238890
GENERAL INFORMATION:
APPLICANT: BOIWE, IYING
APPLICANT: MOYDE, WILLIAM R.
TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
TITLE OF INVENTION: GLYCOPROTEIN HORMONE QUARTET
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW, suite 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,288
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/282,357
FILING DATE:
APPLICATION NUMBER: 08/853,524
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: 08/199,382
FILING DATE: 18-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 29500-20050.25
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
TELEFAX: 202-887-0763
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 681 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-918-288-31

| | Query Match | 12.5% | Score 48.8 | DB 4 | Length 681 |
|----|-----------------------|---|-------------------|------|---------------------------------|
| | Best Local Similarity | 60.6% | Pred. No. 3.8e-05 | | |
| | Matches | 80 | Conservative | 0 | Mismatches 52; Indels 0; Gaps 0 |
| Qy | 250 | tgtaacctcaaacagagaccnaacaggtgactgtgcaagctgcccacactgtgtccccggagatc | 309 | | |
| | | | | | |
| Db | 471 | tgtactcttcaaggaagactggatattatgamaacagtagagagtgccggctgtgcttcacattgca | 412 | | |
| Qy | 310 | gaaccctctcaaacctatccgtgtgcccactcgctgtgactcgagagcctgtctccactgcc | 369 | | |
| | | | | | |
| Db | 411 | gatttccttgtaatacattaccacccagtgccacccagtgctgctactgtggcgaagtgtagacagccagc | 352 | | |
| Qy | 370 | accacagagatgt | 381 | | |
| | | | | | |
| Db | 351 | agcactgattgt | 340 | | |

RESULT 2
 US-09-282-357-31/C
 Sequence 31. Application US/09282357
 Patent No. 6242580
 GENERAL INFORMATION:
 APPLICANT: BOIME, Irving
 APPLICANT: MOYLE, William R.
 TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
 TITLE OF INVENTION: GLYCOPROTEIN HORMONE QUARTET
 NUMBER OF SEQUENCES: 83
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 2000 Pennsylvania Avenue, NW, suite 5500
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1888
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: Fastseq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/282,357
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/918,288
 FILING DATE: 25 AUG-1997
 APPLICATION NUMBER: 08/853,524
 FILING DATE: 09-MAY-1997
 APPLICATION NUMBER: 08/199,382
 FILING DATE: 18-FEB-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Murashige, Kate H
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 29500-20050.25
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-887-1500
 TELEFAX: 202-887-0763
 TELEX:
 INFORMATION FOR SEQ ID NO: 31:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 681 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-09-282-357-31

| | | | | |
|-----------------------|-------|--|------|----------------------------------|
| Query Match | 12.5% | Score 48.8 | DB 4 | Length 681 |
| Best Local Similarity | 60.6% | Pred. NO. 3.8e-05 | | |
| Matches | 80 | Conservative | 0 | Mismatches 52; Indels 0; Gaps 0; |
| QY | 250 | tgtaactacaagagacaaacagtgatgcgaagctgcgccaactgtgtcccccgaagtc | 309 | |
| | | | | |

| | | | |
|----|-----|---|-----|
| Db | 471 | TGTACCTTCAGGAAACTGTTATGAAAACATGAGAGTGCGCCGCTGTGCTCACCATGTCA | 412 |
| Qy | 310 | gaccacctctaacacataaccgctggcagtcgcgtctgacgtggagagcgtctccatgyc | 369 |
| | | | |
| Db | 411 | GATTCTCTGTATACATACCCAGTGGCCACCACGATCTCACTGTGGCAAGTGTGACAGGCAC | 352 |
| Qy | 370 | accacgagagtg | 381 |
| | | | |
| Db | 351 | AGCACTGATGT | 340 |

```

1      RESULT      3
2      US-08-918-288-28/c
3      : Sequence 28, Application US/08918288
4      : Patent No. 6238890
5      : GENERAL INFORMATION:
6      : APPLICANT: BOIME, Irving
7      : APPLICANT: MOYLE, William R.
8      : TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
9      : TITLE OF INVENTION: GLYCOPROTEIN HORMONE QUARTER
10     : NUMBER OF SEQUENCES: 83
11     : CORRESPONDENCE ADDRESS:
12     : ADDRESSEE: MORRISON & FOERSTER
13     : STREET: 2000 Pennsylvania Avenue, NW, suite 5500
14     : CITY: Washington
15     : STATE: DC
16     : COUNTRY: USA
17     : ZIP: 20006-1888
18     : COMPUTER READABLE FORM:
19     : MEDIUM TYPE: Diskette
20     : COMPUTER: IBM Compatible
21     : OPERATING SYSTEM: DOS
22     : SOFTWARE: FastSeq for Windows Version 2.0
23     : CURRENT APPLICATION DATA:
24     : APPLICATION NUMBER: US/08/918.288
25     : FILING DATE:
26     : CLASSIFICATION:
27     : PRIOR APPLICATION DATA:
28     : APPLICATION NUMBER: 09/282.357
29     : FILING DATE:
30     : APPLICATION NUMBER: 08/853.524
31     : FILING DATE: 09-MAY-1997
32     : APPLICATION NUMBER: 08/199.382
33     : FILING DATE: 18-FEB-1994
34     : ATTORNEY/AGENT INFORMATION:
35     : NAME: Murashige, Kate H
36     : REGISTRATION NUMBER: 29,959
37     : REFERENCE/DOCKET NUMBER: 29500-20050.25
38     : TELECOMMUNICATION INFORMATION:
39     : TELEPHONE: 202-887-1500
40     : TELEFAX: 202-887-0763
41     : TELEX:
42     : INFORMATION FOR SEQ ID NO: 28:
43     : SEQUENCE CHARACTERISTICS:
44     : LENGTH: 693 base pairs
45     : TYPE: nucleic acid
46     : STRANDEDNESS: double
47     : TOPOLOGY: linear
48     :
49     : US-08-918-288-28

```

[illegible]


```

; TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
; TITLE OF INVENTION: GLYCOPROTEIN HORMONE QUARTET
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/282,357
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/918,288
; FILING DATE: 25 AUG-1997
; APPLICATION NUMBER: 08/853,524
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: 08/199,382
; FILING DATE: 18-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 29500-20050.25
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-887-0763
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 702 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-282-357-13

Query Match      12.5%; Score 48.8; DB 4; Length 702;
Best Local Similarity 60.6%; Pred. No. 3.9e-05;
Matches 80; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 250 ttactacaagagacaacagtgactgtcaagctgcccgaactgtgccccggagatc 309
DB 492 tttaccttcaaggaactggtatattgaacagtgagagtcctccggctgtcctcaccatgca 433
QY 310 gaccctctacacatcccgltggccatccgctgtgactcgagagcctgctccactgcc 369
DB 432 gattctcttatatacattaccacagtgccacccagtgctcactgtggcagatgtgacacagcag 373
QY 370 accaaggagtg 381
DB 372 agcactgattgt 361
```

```

RESULT 7
US-08-918-288-29
; Sequence 29, Application US/08918288
; Patent No. 6238890
; GENERAL INFORMATION:
; APPLICANT: BOIME, Irving
; APPLICANT: MOYLE, William R.
; TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY:
; STATE:
```

```

; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/918,288
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/282,357
; FILING DATE:
; APPLICATION NUMBER: 08/853,524
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: 08/199,382
; FILING DATE: 18-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 29500-20050.25
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-887-0763
; TELEX:
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 707 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 33...698
; OTHER INFORMATION:
; US-08-918-288-29
```

```

Query Match      12.5%; Score 48.8; DB 4; Length 707;
Best Local Similarity 60.6%; Pred. No. 3.9e-05;
Matches 80; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
```

```

QY 250 ttactacaagagacaacagtgactgtcaagctgcccgaactgtgccccggagatc 309
DB 237 tttaccttcaaggaactggtatattgaacagtgagagtcctccggctgtcctcaccatgca 296
QY 310 gaccctctacacatcccgltggccatccgctgtgactcgagagcctgctccactgcc 369
DB 297 gattctcttatatacattaccacagtgccacccagtgctcactgtggcagatgtgacacagcag 356
QY 370 accaaggagtg 381
DB 357 agcactgattgt 368
```

```

RESULT 8
US-09-282-357-29
; Sequence 29, Application US/09282357
; Patent No. 6242360
; GENERAL INFORMATION:
; APPLICANT: BOIME, Irving
; APPLICANT: MOYLE, William R.
; TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
```

```

? COUNTRY: USA
? ZIP: 20006-1888
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/282,357
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/918,288
? FILING DATE: 25 AUG-1997
? APPLICATION NUMBER: 08/853,524
? FILING DATE: 09-MAY-1997
? APPLICATION NUMBER: 08/199,382
? FILING DATE: 18-FEB-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Murashige, Kate H
? REGISTRATION NUMBER: 29,959
? REFERENCE/DOCKET NUMBER: 29500-20050.25
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-887-1500
? TELEFAX: 202-887-0763
? TELEX:
? INFORMATION FOR SEQ ID NO: 29:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 707 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? FEATURE:
? NAME/KEY: Coding Sequence
? LOCATION: 33...698
? OTHER INFORMATION:
?
US-09-282-357-29

Query Match          12.5%; Score 48.8; DB 4; Length 707;
Best Local Similarity 60.6%; Pred. No.3.9e-05;
Matches 80; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 250 tgaactacaagagacaaacagagtgactgcaagctgcccgaactgtgcccggagagtc 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 237 tgaactctcaagaaactgataatgaacagtgagagtgcccgctgtgctcacacatgca 296
QY 310 gaccctctcaacactatcccgtyggccatccgcgtgtgactgagcgagcctctccactgcc 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 297 gattccttgatatacattaccagtgccaccacagtgctcactgtggcaagtgtgacacggac 356
QY 370 accacgagagtgt 381
    ||||| |||||
DB 357 agcactgattgt 368

RESULT 9
US-08-918-288-26
; Sequence 26, Application US/08918288
; Patent No. 6238890
; GENERAL INFORMATION:
; APPLICANT: BOIME, Irving
; APPLICANT: MOYLE, William R.
; TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
; TITLE OF INVENTION: GLYCOPROTEIN HORMONE QUARTER
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
```

```

? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/918,288
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 09/282,357
? FILING DATE:
? APPLICATION NUMBER: 08/853,524
? FILING DATE: 09-MAY-1997
? APPLICATION NUMBER: 08/199,382
? FILING DATE: 18-FEB-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Murashige, Kate H
? REGISTRATION NUMBER: 29,959
? REFERENCE/DOCKET NUMBER: 29500-20050.25
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-887-1500
? TELEFAX: 202-887-0763
? TELEX:
? INFORMATION FOR SEQ ID NO: 26:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 719 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? FEATURE:
? NAME/KEY: Coding Sequence
? LOCATION: 33...700
? OTHER INFORMATION:
?
US-08-918-288-26

Query Match          12.5%; Score 48.8; DB 4; Length 719;
Best Local Similarity 60.6%; Pred. No.3.9e-05;
Matches 80; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 250 tgaactacaagagacaaacagagtgactgcaagctgcccgaactgtgcccggagagtc 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 237 tgaactctcaagaaactgataatgaacagtgagagtgcccgctgtgctcacacatgca 296
QY 310 gaccctctcaacactatcccgtyggccatccgcgtgtgactgagcgagcctctccactgcc 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 297 gattccttgatatacattaccagtgccaccacagtgctcactgtggcaagtgtgacacggac 356
QY 370 accacgagagtgt 381
    ||||| |||||
DB 357 agcactgattgt 368

RESULT 10
US-09-282-357-26
; Sequence 26, Application US/09282357
; Patent No. 6242580
; GENERAL INFORMATION:
; APPLICANT: BOIME, Irving
; APPLICANT: MOYLE, William R.
; TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
; TITLE OF INVENTION: GLYCOPROTEIN HORMONE QUARTER
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```

;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/282,357
;   FILING DATE:
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/918,288
;     FILING DATE: 25 AUG-1997
;     APPLICATION NUMBER: 08/853,524
;     FILING DATE: 09-MAY-1997
;     APPLICATION NUMBER: 08/199,382
;     FILING DATE: 18-FEB-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Murashige, Kate H
;     REGISTRATION NUMBER: 29,959
;     REFERENCE/DOCKET NUMBER: 29500-20050.25
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-887-1500
;     TELEFAX: 202-887-0763
;   TELEX:
;   INFORMATION FOR SEQ ID NO: 26:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 719 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;   FEATURE:
;     NAME/KEY: Coding Sequence
;     LOCATION: 33...700
;     OTHER INFORMATION:
;
; US-09-282-357-26
;
Query Match      12.5%; Score 48.8; DB 4; Length 719;
Best Local Similarity 60.6%; Pred. No. 3.9e-05;
Matches 80; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
;
QY 250 tctactacaagacaaacagtgactgtcaagctgtgcccaactgtgccccggagtc 309
    ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 237 tctaccttcacgaagactgtatgatgaacagtgagagtcgccggctgtcctcaccatgca 296
;
QY 310 gacccttctaacctatccgctggccatccgctgtgactgtggagccctgcctccactgcc 369
    || ||||| ||| ||| ||| ||||| ||| ||||| ||| ||| ||| ||| ||| |||
DB 297 gatctcctgtatatacattccacagtgccacccagctgtcactgtggcgaagtgtacacgacgac 356
;
QY 370 accacggagtgct 381
    | ||| ||| |||
DB 357 agcactgattgt 368
;
;
RESULT 11
; US-08-918-288-11
; Sequence 11, Application US/08918288
; Patent No. 6238990
; GENERAL INFORMATION:
;   APPLICANT: BOIME, Irving
;   APPLICANT: MOYLE, William R.
;   TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
;   TITLE OF INVENTION: GLYCOPROTEIN HORMONE QUARTET
;   NUMBER OF SEQUENCES: 83
;   CORRESPONDENCE ADDRESS:
;     ADDRESSER: MORRISON & FOERSTER
;     STREET: 2000 Pennsylvania Avenue, NW, suite 5500
;     CITY: Washington
;     STATE: DC
;     COUNTRY: USA
;     ZIP: 20006-1888
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     CURRENT APPLICATION DATA:

```

```

;
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/918,288
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 09/282,357
;     FILING DATE:
;     APPLICATION NUMBER: 08/853,524
;     FILING DATE: 09-MAY-1997
;     APPLICATION NUMBER: 08/199,382
;     FILING DATE: 18-FEB-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Murashige, Kate H
;     REGISTRATION NUMBER: 29,959
;     REFERENCE/DOCKET NUMBER: 29500-20050.25
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-887-1500
;     TELEFAX: 202-887-0763
;   TELEX:
;   INFORMATION FOR SEQ ID NO: 11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 728 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;   FEATURE:
;     NAME/KEY: Coding Sequence
;     LOCATION: 33...719
;     OTHER INFORMATION:
;
; US-08-918-288-11
;
Query Match      12.5%; Score 48.8; DB 4; Length 728;
Best Local Similarity 60.6%; Pred. No. 3.9e-05;
Matches 80; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
;
QY 250 tctactacaagacaaacagtgactgtcaagctgtgcccaactgtgccccggagtc 309
    ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 237 tctaccttcacgaagactgtatgatgaacagtgagagtcgccggctgtcctcaccatgca 296
;
QY 310 gacccttctaacctatccgctggccatccgctgtgactgtggagccctgcctccactgcc 369
    || ||||| ||| ||| ||| ||||| ||| ||||| ||| ||| ||| ||| ||| |||
DB 297 gatctcctgtatatacattccacagtgccacccagctgtcactgtggcgaagtgtacacgacgac 356
;
QY 370 accacggagtgct 381
    | ||| ||| |||
DB 357 agcactgattgt 368
;
;
RESULT 12
; US-09-282-357-11
; Sequence 11, Application US/09282357
; Patent No. 6242580
; GENERAL INFORMATION:
;   APPLICANT: BOIME, Irving
;   APPLICANT: MOYLE, William R.
;   TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
;   TITLE OF INVENTION: GLYCOPROTEIN HORMONE QUARTET
;   NUMBER OF SEQUENCES: 83
;   CORRESPONDENCE ADDRESS:
;     ADDRESSER: MORRISON & FOERSTER
;     STREET: 2000 Pennsylvania Avenue, NW, suite 5500
;     CITY: Washington
;     STATE: DC
;     COUNTRY: USA
;     ZIP: 20006-1888
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FASTSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:

```

```

? APPLICATION NUMBER: US/09/282,357
? FILING DATE:
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/918,288
? FILING DATE: 25 AUG-1997
? APPLICATION NUMBER: 08/853,524
? FILING DATE: 09-MAY-1997
? APPLICATION NUMBER: 08/199,382
? FILING DATE: 18-FEB-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Murashige, Kate H
? REGISTRATION NUMBER: 29,959
? REFERENCE/DOCKET NUMBER: 29500-20050.25
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-887-1500
? TELEFAX: 202-887-0763
? TELEX:
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 728 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? FEATURE:
? NAME/KEY: Coding Sequence
? LOCATION: 33...719
? OTHER INFORMATION:
? US-09-282-357-11

Query Match      12.5%; Score 48.8; DB 4; Length 728;
Best Local Similarity 60.6%; Pred. No. 3.9e-05;
Matches 80; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 250 tgaactacaagagacaaacaggtgactgtcaagctgcccacactgtgccccggagtc 309
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 237 tgaaccttcaggaactggtatgtgaaacagtgagtgcccgctgtctacacatgca 296
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 gaccctctacacatcccggtgacatccgctgtgactgagcgagccttcacatgcc 369
    || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 297 gattccttgatgacatgacacagtgccacccacagtgctgctgacgaagtgtgacagcgac 356
    || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 370 accacggagtgct 381
    || ||| ||| |||
DB 357 agcactgattgt 368
    || ||| ||| |||

RESULT 13
US-08-239-256-1
? Sequence 1, Application US/08239256
? Patent No. 5585345
? GENERAL INFORMATION:
? APPLICANT: BOIME, IRVING
? APPLICANT: MATYUK, MARTIN M.
? APPLICANT: KEENE, JEFFREY L.
? TITLE OF INVENTION: CTP EXTENDED FORM OF LH
? NUMBER OF SEQUENCES: 24
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: MORRISON & FOERSTER
? STREET: 2000 Pennsylvania Ave. N.W.
? CITY: Washington, D.C.
? COUNTRY: USA
? ZIP: 20006-1812
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/239,256
? FILING DATE: 06-MAY-1994
? CLASSIFICATION: 435
```

```

? ATTORNEY/AGENT INFORMATION:
? NAME: MURASHIGE, KATE H.
? REGISTRATION NUMBER: 29,959
? REFERENCE/DOCKET NUMBER: 29500-20030.12
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 887-1500
? TELEFAX: (202) 887-0763
? TELEX: 90-4030
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1006 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 238..396
? FEATURE:
? NAME/KEY: mat_peptide
? LOCATION: 292
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 461..686
? US-08-239-256-1

Query Match      12.5%; Score 48.8; DB 1; Length 1006;
Best Local Similarity 60.6%; Pred. No. 4.3e-05;
Matches 80; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 250 tgaactacaagagacaaacaggtgactgtcaagctgcccacactgtgccccggagtc 309
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 506 tgaaccttcaggaactggtatgtaacagtgagtgcccgctgtctacacatgca 565
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 gaccctctacacatcccggtgacatccgctgtgactgagcgagccttcacatgcc 369
    || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 566 gattccttgatgacatgacacagtgccacccacagtgctgctgacgaagtgtgacagcgac 625
    || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 370 accacggagtgct 381
    || ||| ||| |||
DB 626 agcactgattgt 637
    || ||| ||| |||

RESULT 14
US-08-918-288-10/c
? Sequence 10, Application US/08918288
? Patent No. 6238890
? GENERAL INFORMATION:
? APPLICANT: BOIME, IRVING
? APPLICANT: MOYLE, WILLIAM R.
? TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE
? TITLE OF INVENTION: GLYCOPROTEIN HORMONE QUARTET
? NUMBER OF SEQUENCES: 83
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: MORRISON & FOERSTER
? STREET: 2000 Pennsylvania Avenue, NW, suite 5500
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20006-1888
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/918,288
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 09/282,357
? FILING DATE:
? APPLICATION NUMBER: 08/853,524
```


THIS PAGE BLANK (USPTO)